ANALYSIS OF SOLUTION-PHASE MACROMOLECULAR MATERIALS BY DIFFUSION NMR

Edward Eden

Primary supervisor: Prof A I Cooper

Secondary supervisor: Prof D J Adams

Thesis submitted in accordance with the requirements of the

University of Liverpool for the degree of Doctor of Philosophy

“They shot the tail,

but they missed the fox.”

ROALD DAHL

Abstract

Molecules such as covalent cages can adopt several shapes, in which the ratio of starting materials is the same, but the number of starting materials, and the shape of the resulting molecule is different. However, determining this in the absence of a crystal structure can be challenging.

Pulsed field gradient (PFG) NMR has been used for two decades to characterise large macromolecules in solution, but it is still difficult to determine precise structural information, because of the rotational-averaging experienced in experimental measurements.

Here, we develop experimental techniques for collecting PFG-NMR data that break this barrier, and allow characterisation of several useful molecular descriptors. By measuring the diffusion coefficients of molecules in a range of solvents, incurvate surfaces are probed to map the outer surface of nanometre-sized molecular species. This technique allows details about the geometrical shape of covalent cages to be determined without the need for isolation.

Furthermore, we compare experimental PFG-NMR data to structures produced by computational modelling and produce a new molecular descriptor, ρ_r, which describes the isotropy of covalent cages. This descriptor is used to determine the quality of agreement between proposed structures and experimental PFG-NMR data.

In analysing polymers, we develop a new mathematical model for determining the molar-mass dispersity ($Đ_M$) by PFG-NMR. We find a single parameter is sufficient to determine the dispersity of a system, which eliminates the need for data modelling and enhances the reliability of analysis. We hope this will make the technique more

accessible to polymer scientists, and will help test the validity of molar-mass dispersity measurements made by other means.

Finally, we synthesise five novel dodecaamide cages, which contain functional groups that offer the opportunity to extend functionality beyond the cage *via* further reaction. We take significant steps towards producing singly functionalised species, which could be incorporated into polymeric materials for the development of robust membranes and coatings.

Acknowledgements

My primary supervisor, Andy Cooper, deserves thanks for years of advice, sarcasm and friendship. I am indebted to him for his support and encouragement. Over four years, he has fired and re-hired me seven times, always citing the quality of my jokes (surely the sort of stuff that drives you to improve). I will always be grateful for the autonomy and respect that he has ladled out, and I am always astounded at the size of the ladle.

My secondary supervisor Dave Adams and Michael Briggs have contributed hugely to the work that has gone into writing this thesis. Their day-to-day advice has helped with synthetic problems and project development, and they have spent a great amount of time reading drafts.

Becky Greenaway and Anna Slater warrant particular thanks for their help as colleagues and their kindness as friends. As a source of coffee, cake and common sense, they might just be the best.

My fellow group members, especially Mike Bennison, Ben Slater, Meera Vijayaraghavan and Steven Robinson (honorary), have made my time in Liverpool tremendously fun, and they have all helped in one way or another. Thank you, all.

I'd like to thank Stephen Moss, Moya McCarron and Jean Ellis for their help with mass spectrometry. In NMR, Jon Iggo, Matthew Wallace, Konstantin Luzyanin and Tony Ellis have always been hugely accommodating with their advice and time. Matthew deserves special thanks for showing me how to run diffusion NMR. Sean Higgins has run a great number of GPCs over four years. George Miller was very patient in attempting to get CHN data for **CC-pentane**. Because of the nature of this

thesis, a huge number of people have helped to some degree and specific author contributions have been listed in a special Declaration of Contributions (page iv).

Thank you Lisa, who has supported and encouraged me, frequently by finding jobs I can do in the garden, but only once I've finished a Chapter. Quinton and Alex, who find joy in the most ridiculous places, have taught me to find joy there too. Finally, thank you to my parents, Lesley and Nick, and brother Rich, who look on tempests and are never shaken. They have cooled my rants and supported me for 29 years, never mind these four. They are the very, very best people and I owe them everything.

List of Publications

K. E. Jelfs, **E. G. B. Eden**, J. L. Culshaw, S. Shakespeare, E. O. Pyzer-Knapp, H. P. G. Thompson, J. Bacsa, G. M. Day, D. J. Adams, and A. I. Cooper, *Journal of the American Chemical Society*, **2013**, *135*, 9307–9310.

C. Colquhoun, E. R. Draper, **E. G. B. Eden,** B. N. Cattoz, K. L. Morris, L Chen, T. O. McDonald, A. E. Terry, P. C. Griffiths, L. C. Serpell and Dave J. Adams, *Nanoscale*, **2014**, *6*, 13719–13725.

E. R. Draper, **E. G. B. Eden**, T. O. McDonald, D. J. Adams, *Nature Chemistry*, **2015**, *7*, 848–852.

R. L. Greenaway, D. Holden, **E. G. B. Eden**, A. Stephenson, C. W. Yong, M. J. Bennison, T. Hasell, M. E. Briggs, S. L. James and A. I. Cooper, *Chemical Science*, Understanding Gas Capacity, Guest Selectivity and Diffusion in Porous Liquids, DOI: 10.1039/C6SC05196K.

E. G. B. Eden, T. R. Congdon, B. Martyn, M. Barrow, T. Ratvijitvech, M. E Briggs, M. I. Gibson, A. I. Cooper, and D. J. Adams, *Polymer Chem.*, A Robust and Convenient Method for Determining Dispersity by ^{1}H NMR Spectroscopy, Manuscript in preparation.

Declaration of contributions

This thesis focusses on the analysis of functional molecular materials and, frequently, these materials have been provided by collaborators and colleagues. These details are highlighted throughout the thesis, but have been summarised here for clarity.

Molecular modelling: with the exception of Chapter 2, Section 2.2.4, all computational modelling was performed by the Jelfs group at the Imperial College London. Structures of **CC-propane**, **CC-butane**, **CC-pentane**, **CC-hexane**, **CC-heptane**, **CC-octane** and **CC-nonane** were produced by Kim Jelfs. The remaining structures in Chapters 2 and 4 were generated by Valentina Santolini. The molecular modelling in Chapter 2, Section 2.2.4, concerning porous liquids, was produced by Daniel Holden.

Synthetic contributions: Several cages were supplied by colleagues at the University of Liverpool: the **ether-cage** was synthesised by Heather Armstrong; **4OH-CC3** and **Ico-cage** were synthesised by Baiyang Teng, and the **Astetrapod** was synthesised by Xiaofeng Wu. Samples of **FT-CC3** and **RCC7** were provided by Ming Liu in addition to the material we produced ourselves. All cages in Chapter 2, Section 2.2.3 were synthesised by Becky Greenaway, and robotic screening was conducted by Ben Alston. All non-spectroscopic analysis was conducted by members of the project team. The scrambled cages produced in Chapter 2, Section 2.2.4 were synthesised by Becky Greenaway, and all non-spectroscopic analysis was performed by members of the

project team. All optimisation and gas uptake tests were performed by Becky Greenaway. In Chapter 3, polymers **PS1–3** were synthesised by Tanchanok Ratvijitvech; **MPC1–2** were synthesised by Michael Barrow; **PVA1**, **PVP1** and **PVAP1** were synthesised by Tom Congdon at the University of Warwick; and **PHEAA1–3** were synthesised by Ben Martyn at the University of Warwick. Matthew Gibson lead the Warwick team.

Abbreviations

COF	Covalent organic framework
COSY	Correlation spectroscopy
CSA	Connolly Surface Area
CSD	Cambridge Structural Database
DCC	Dynamic covalent chemistry
DCM	Dichloromethane
DFT	Density functional theory
DMF	*N*,*N*′-dimethylformamide
DIBALH	Diisobutylaluminium hydride
DOSY	Diffusion ordered spectroscopy
FT-IR	Fourier-transform infra-red
HCP	Hexachloropropene
HMBC	Heteronuclear multiple-bond correlation
HNFX	3,3,7,7-Tetrakis(difluoramino)octahydro-1,5-dinitro-1,5-diazocine
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single-quantum correlation
Hz	Hertz (per second)
HT	High-throughput
ICL	Imperial College London
ILT	Inverse Laplace transformation
IPA	Isopropylalcohol
IUPAC	International Union of Pure and Applied Chemistry
MC	Monte Carlo

MD	Molecular dynamics
MOF	Metal-organic framework
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
UV	Ultraviolet
PAF	Porous aromatic framework
PFG	Pulsed field gradient
PIM	Polymers of intrinsic microporosity
PPN	Porous polymer network
Ppm	Parts per million
Ppt	Precipitate
RF	Radio frequency
RMS	Root mean squared
SCXRD	Single crystal X-ray crystallography
SIC	Small imine control
STE	Stimulated echo
TDMB	1-(*tert*-butyl)-3,5-dimethylbenzene
TFB	Triformylbenzene
TLC	Thin layer chromatography
TMS	Tetramethylsilane
TPP	Tris(phenylenedioxy)cyclo-triphosphazene
TTEB	3,3’,4,4’-Tetra-kis(trimethylsilylethynyl)biphenyl
VdW	Van der Waals
ZIF	Zeolitic imidazolate framework

Chapter 1. Table of Contents

Chapter 1. General Introduction

Measuring the properties of functional materials is fundamental to determining their likely function. The pore size of molecular sorbents, for example, dictates the possible guests and applications of a material.[1] Furthermore, the processing of materials has become increasingly important as factors such as scalability, capital costs, and energy efficiency determine the viability of new materials.[2–4] As the structure of sorbents has changed over the past 50 years, so too have the methods of characterising them.[5–12] Zeolites, porous polymers and many framework solids are insoluble,[13,14] which limits both subsequent processing and, sometimes, function. Soluble porous materials, such as porous molecular materials[15] and soluble porous polymers,[16] can be processed in solution. In this Chapter, we discuss the advantages and disadvantages of each structural type.

Section 1.2 summarises extended networks, which were discovered early in the development of porous materials as a result of the natural abundance of zeolites. Zeolites, which occur naturally, were first reported in 1765 by Axel Fredrik Cronstedt,[17] but their development as adsorbents was not developed until the 1950s, when Union Carbide monetised gas and liquid separations.[18] Since then, examples of separations using activated carbons[19,20] and metal-organic frameworks (MOFs),[21,22] have demonstrated the utility in investigating microporous framework solids as sorbents. Section 1.3 introduces porous organic cages, which were discovered in 2008 by Cooper[15] and Mastalerz.[23] These intrinsically porous materials are distinguished from extrinsically porous molecular materials, where the porosity is derived from poor three-dimensional packing. Methods of characterising both intrinsically and extrinsically porous materials, both in the solid state, and in solution, are discussed.

Finally, Section 1.4 discusses porous polymeric materials. Although many of these are insoluble, porous polymers can benefit from both solution processability and porosity. Characterisation of these materials *in situ* is often difficult, due to high viscosities[24,25] and internal temperatures during synthesis,[26,27] and we propose a method for measuring some molecular descriptors without the need for isolation.

1.1 Porous framework solids

Microporous frameworks have enjoyed significant research interest over the last two decades, as a result of applications developed in separations,[19–22] gas storage,[28-30] and catalysis.[13,31] Much of the development within this field was, until recently, limited to inorganic and hybrid organic-inorganic materials such as zeolites,[32] metal-organic frameworks (MOFs),[33] covalent organic frameworks (COFs),[34] and porous organic polymers.[35] All four classes of material can be characterised as extended networks (Figure 1.1), although porous polymers can be distinguished from other extended networks as their long range connectivity is derived from entanglement of long fibres.

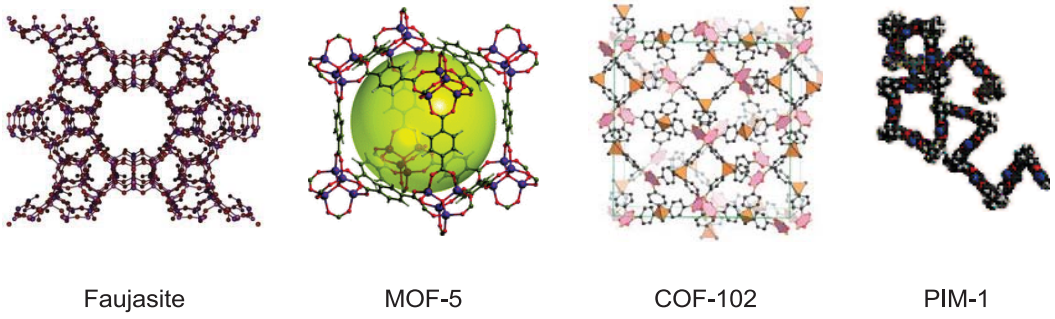


Figure 1.1. Representative examples of extended three-dimensional structures. The representation of Faujasite has been modified from reference 36. The representation of MOF-5 has been reproduced from reference 37; PIM-1 has been reproduced from reference 38. The representation of PIM-1 has been reproduced from reference 39.

These materials are characterised by a thermal stability that is much higher than that of porous molecular materials,[40–44] and many have a rigid topology.[11,16,45–50] More recently, flexible MOFs have been synthesised,[12,51–53] although molecular mobility is still limited by covalent bonding between framework components. Framework materials are often characterised by X-ray diffraction[11] and sorption analysis.[6,10,54] These give details of the underlying pore structure, which can be further investigated through spectroscopy of probe-molecules using absorption spectroscopies,[9] and NMR.[55,56] Metal sites, flexibility, and assembly can be studied using Raman spectroscopy,[5,57,58] and morphology can be studied using electron microscopy.[7,8]

1.2 Porous molecular materials

Porous molecular materials differ from extended networks in the construction of their long-range three-dimensional structure (Figure 1.2). Porous molecular materials are defined here as those assembled through interactions between small molecules, such as covalent cages[15,59–62] and extrinsically porous molecules.[63,64] In place of the strong, directional bonding present in framework materials, molecular solids are held together by intermolecular interactions, which are generally weaker.[65–67] This can result in molecular mobility,[39] polymorphism,[68–70] and can complicate crystal design.[71] However, porous molecular materials have advantages as alternatives to framework materials as a result of their solution-processability.[72] This is crucial in areas such as membranes and coatings, where functional materials are often cast from solution.[73,74]

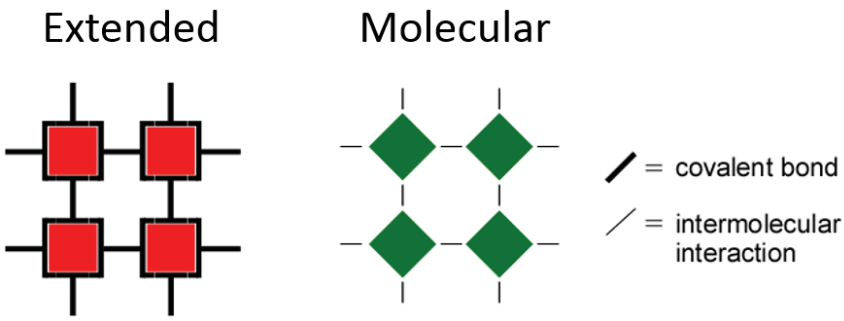


Figure 1.2. The extended structure of porous networks is composed of strong, covalent bonds. Conversely, the crystal structure of porous molecular materials is defined by intermolecular interactions.

Weak non-covalent interactions that constitute the extended structure of porous molecular materials are less directional than those found in zeolites and MOFs. This makes rational design more challenging. Thus, development of porous molecular materials has post-dated MOFs by about 10 years, and high surface areas have only recently been achieved.[75,76]

1.2.1 The history of porous organic materials

Porous molecular materials are often formed by removing guests such as solvent molecules from an inclusion compound, which can be accomplished under reduced pressure (Figure 1.3). In many cases, structures are observed to collapse entirely, such that no permanent pores remain within the structure. Frequently, solvent-exchange and slow evaporation are required to remove guests with enough care to leave the structure intact.[63,77]

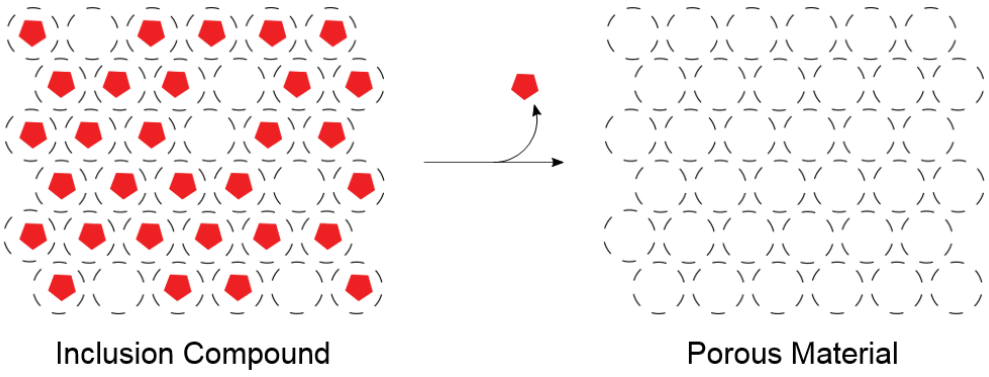


Figure 1.3. Simplified representation of a porous material, synthesised by removing guest molecules (red) from an inclusion compound.

Since the mid-1990s, access to porous molecular materials has increased significantly, as improvements in the design of molecular architecture[78] and crystal engineering[79] have widened the scope of material design.

1.2.2 Gossypol – the first porous molecular material

In 1994, the desolvated crystal structure of Gossypol was demonstrated to maintain one-dimensional pores large enough to incorporate several halogeno-methanes into one-dimensional channels (Figure 1.4).[80] Several polymorphs were synthesised by crystallising Gossypol from different halogeno-methanes. The polymorphs obtained varied in cell parameter, and channel radius. In the dichloromethane polymorph, a channel diameter of 5.6 Å was reported,[80] non-inclusive of the Van der Waals radius of atoms that surrounded the cavity. This reduced by 4–9 % on desolvation.

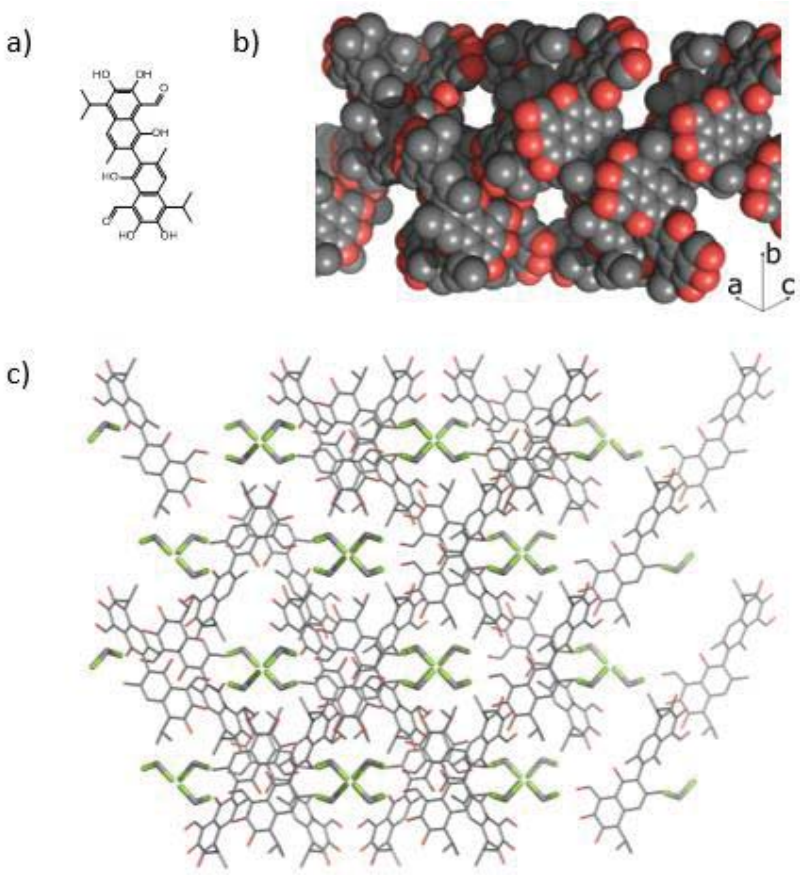


Figure 1.4. (a) The chemical structure of Gossypol; (b) a space-filling representation of the three-dimensional structure obtained from reference 80. Solvent molecules have been removed to show the pore structure; (c) Crystal structure of Gossypol obtained from reference 80. DCM and Gossypol are drawn as capped stick representations. Dichloromethane has been shown as a capped stick representation.

1.2.3 First generation porous molecular materials

Subsequently, an increasing number of porous materials were reported towards the turn of the century. By 1999, the crystal structure of 3,3,7,7-tetrakis(difluoramino)octahydro-1,5-dinitro-1,5-diazocin (HNFX)[81] had been shown to contain pores that survived desolvation, and Dianin's compound was demonstrated to maintain its structure on removal of guests (Figure 1.5).[82]

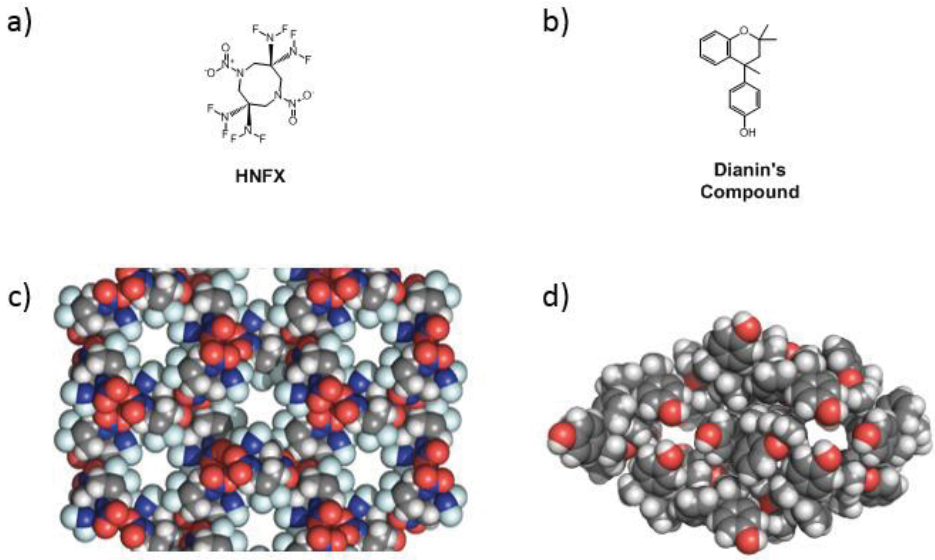


Figure 1.5. Chemical structures of (a) HFNX and (b) Dianin's compound; Crystal structures of (c) HFNX and (d) Dianin's compound, using a space-filling representation to show the one-dimensional pores.

In 2000, Sozanni *et al*. synthesised tris(phenylenedioxy)cyclo-triphosphazene (TPP, Figure 1.6a).[83] This represented the first porous molecular material in which only weak non-covalent interactions determined the extended structure. In contrast to HFNX and Dianin's compound, no hydrogen bonding is present to stabilise the structure (Figure 1.6b). Laser-polarised ^{129}Xe NMR spectroscopy demonstrated that pore channels were almost cylindrical, an effect which was produced by the vertical faces of aromatic rings that surround each channel. In 2005, sorption analysis demonstrated that crystals of TPP could be used for xenon and nitrogen storage,[84,85] and in 2006, the Langmuir surface area of TPP was reported to be 240 m^2g^{-1}.[85]

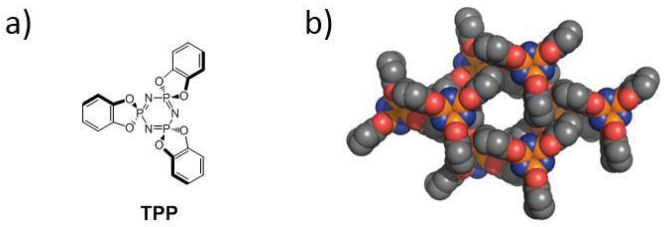


Figure 1.6. (a) The chemical structure of TPP; (b) The crystal structure of TPP, shown using a space-filling representation down the channel axis.

Since the early 2000s, obtainable surface areas have risen significantly,[59,75,76,86] post-dating the rise observed in MOFs ten years before (Figure 1.7).

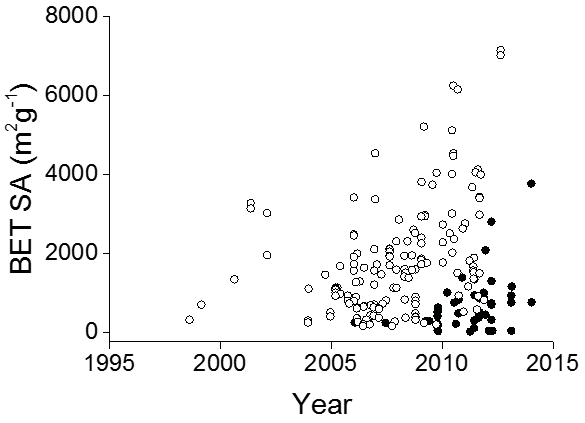


Figure 1.7. The surface areas of MOFs (open circles) and porous molecular materials (closed circles) from 1995 to 2015.

More recently, researchers have focussed on producing molecules with intrinsic porosity,[15,59,61,76,78,86] which is defined as the presence of pores intrinsic to the structure of the molecule, rather than being generated by crystal packing (Figure 1.8). Covalent cages[15] are often constructed as rigid molecular capsules, with an internal pore that is accessible *via* windows in the structure. These are differentiated from metal organic polyhedral in their construction, with covalent cages being synthesised

from organic atoms such as carbon, hydrogen and nitrogen. More recently, boron has been incorporated into these structures.[87]

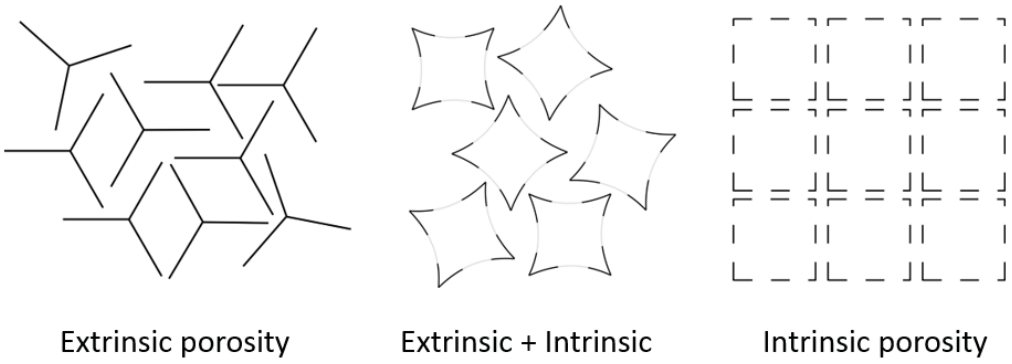


Figure 1.8. A schematic representation of extrinsically and intrinsically porous materials.

Covalent cages are traditionally synthesised by dynamic combinatorial chemistry (DCC, also called dynamic covalent chemistry),[88] a process in which starting materials, products and intermediates exist in a dynamic equilibrium (Figure 1.9).[90-95] DCC facilitates formation of a single product, which represents the most thermodynamically stable configuration of starting materials. This is important in covalent cage synthesis, where the starting materials can adopt a number of geometrical shapes. Frequently, these shapes resemble Platonic or Archimedean solids, such as tetrahedra (Figure 1.9, topology 3), which is referred to as a "topology."

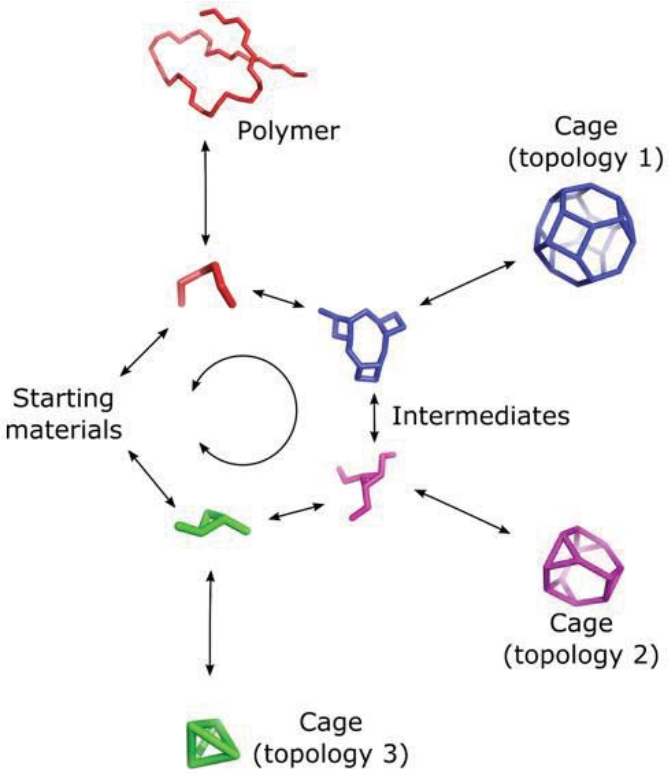


Figure 1.9. Schematic representation of a system under dynamic combinatorial control.

Amidst the growth in reactions that fulfil the criterion of reversibility,[89,90,92,93] cages that contain imine groups have received by far the most attention.[95–97] This is in part, perhaps, because an imine in conjugation with an aromatic system lends itself to predictable geometry, with the C-N=C-C *trans*-imine bond angle being close to zero, both in small molecules, and in covalent cages (Figure 1.10).[98] This has enabled researchers to develop a series of design principles that extend to the construction of covalent cages.[60]

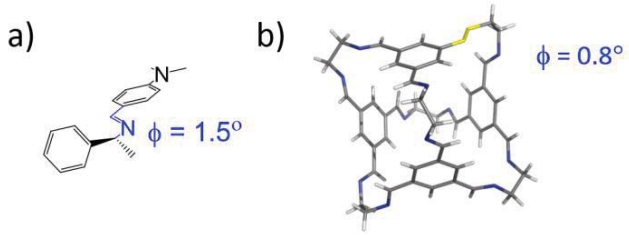


Figure 1.10: (a) Molecular structure of a small molecule imine, with the C-N=C-C *trans*-imine bond angle shown in blue; (b) Crystal structure of CC1[15] shown as a capped stick representation, with one imine bond angle highlighted in yellow. *Trans*-imine bond angle is shown in blue.

1.2.4 Organic cages

Organic cages are often synthesised by mixing multiply-functionalised amine and aldehyde components under conditions that yield a dynamic equilibrium. This can be accomplished by using low concentrations of starting materials, adding one component slowly, and reducing the temperature of the reaction mixture to reduce the rate of reaction.[59]

Several families of covalent cages have been synthesised by Mastalerz and Cooper (Figure 1.11),[15,99–102] both of which demonstrate the ability to exchange geometrically similar starting materials in the synthesis of a series of cages.

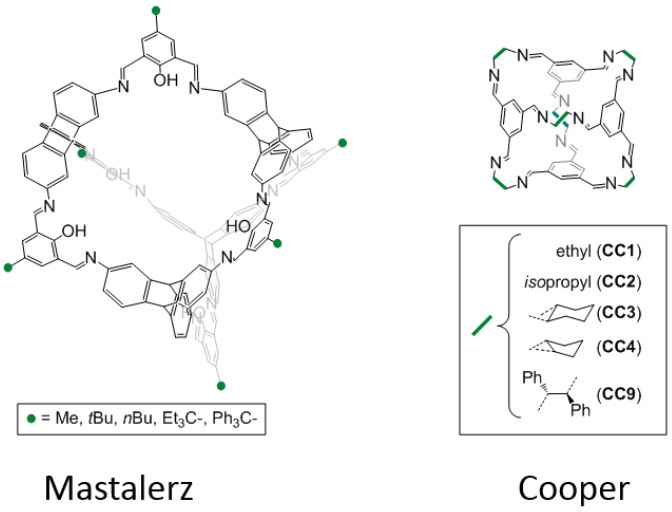


Figure 1.11: Molecular structure of molecular cages synthesised by Mastalerz (left) and Cooper (right).

However, this is not universally the case, as was demonstrated in the synthesis of **CC7** by Jelfs et al.[103] Here, the amine component used in the synthesis if **CC3** is coupled with an aldehyde component that has been extended relative to that used in the synthesis of **CC3** (Figure 1.12). The stoichiometry of covalent cages has been stated as [x+y], where x represents the number of aldehyde molecules, and y represents the number of amine molecules in the final structure.[15] This provides a useful shorthand for referring to cage topologies, as [3+2] cages are prismatic, [4+6] cages are generally tetrahedral, and [8+12] cages are frequently cubic.

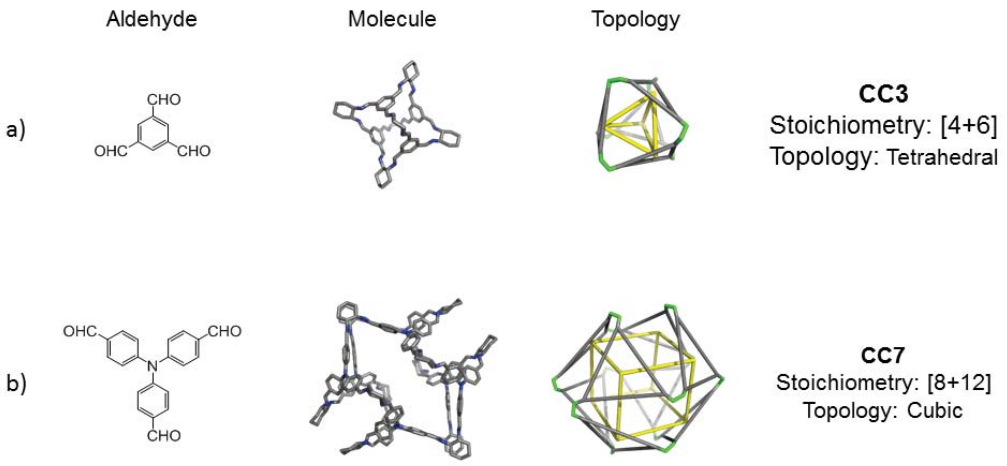


Figure 1.12. (a) Aldehyde starting material, molecular structure and underlying topology of CC3; (b) Aldehyde starting material, molecular structure and underlying topology of CC7. Topologies are shown as a capped stick representation, with the underlying shape highlighted in yellow. For clarity, cyclohexyl groups have been removed, and their positions are highlighted in green. Aldehydes have been simplified as grey triangles. Hydrogens have been omitted for clarity.

Because the underlying structure of a cage can be transformed by making small changes to the molecular structure of its components, it is difficult to predict cage topology from the geometry of its starting materials. This can complicate design in both the synthesis,[104] and subsequent crystal design.[69] In contrast to MOFs, where bonding is highly directional, the non-covalent interactions that hold porous molecular materials together are significantly more complex to predict. In spite of this, crystal packing in porous molecular materials are frequently dominated by Van der Waals forces, and the resultant structures are dictated by the shape of covalent cages.[69,71]

1.2.5 Characterisation of porous molecular materials

Recently, the structures of cages in both solution[104,105] and the solid state[71] have been predicted computationally. This has added to the tool-kit of techniques used to characterise large supramolecules like porous molecular materials, which already

includes X-ray diffraction, mass spectrometry, electron microscopies, sorption analysis, and NMR spectroscopy (Table 1.1).[61,96,106–108]

Acronym/ Name	Description	State
AFM	Atomic force microscopy	Solid
BET	Gas sorption model based on multi-layer adsorption	Solid
FTIR	Fourier transform infrared (spectroscopy)	Both
HPLC	High performance liquid chromatography	Solution
HRTEM	High resolution transmission electron microscopy	Solid
MS	Mass spectrometry	Solution
SANS	Small angle neutron scattering	Both
SEM	Scanning electron microscopy	Solid
STM	Scanning tunnelling microscopy	Solid
MAS-NMR	Magic angle spinning nuclear magnetic resonance (spectroscopy)	Solid
NMR	Nuclear magnetic resonance (spectroscopy)	Both
PFG-NMR	Pulsed field gradient nuclear magnetic resonance	Both
TPD	Temperature programmed desorption	Solid
XRD	X-Ray diffraction.	Solid
XPS	X-ray induced photoelectron spectroscopy	Solid

Table 1.1. Techniques used in the characterisation of porous molecular materials

A more exhaustive list of solid state characterisation methods can be found in reference 13. In support of experimental techniques, crystal engineering, which seeks to isolate contributory factors such as hydrogen bonding,[109,110] π–π interactions,[111] and C–X···π interactions,[112] has contributed to growing confidence in design.[65,66,79,113,114] Furthermore, advances in crystal structure prediction[115–119] have enabled detailed predictions to be generated computationally, and tested experimentally.[71]

The majority of solution-phase techniques used to characterise covalent cages in solution are focussed around establishing the purity of samples produced. High performance liquid chromatography (HPLC) has been used to characterise mixtures of cages,[107] but provides no insight into structural detail. Similarly, mass spectrometry provides details of the mass, but has limited scope for determining the size of a covalent cage. We use mass spectrometry in Chapter 2 to confirm the structures of several novel covalent cages, but go further in developing analytical models that enable structural features to be calculated from pulsed field gradient (PFG) NMR spectroscopic analysis.

1.2.6 Pulsed field gradient NMR in characterising covalent cages

PFG-NMR was developed in the early 1990s,[120] and has been widely used to confirm the structure of macromolecular species.[121] Through the use of magnetic field gradients, PFG-NMR measures the speed with which molecules diffuse through solution. The diffusion coefficient measured, D, is inversely proportional to molecular size.[122] Thus, by measuring the diffusion speed of molecules in solution, PFG-NMR aims to quantitatively measure the size of particles that diffuse through solution (a thorough treatment of the mathematics behind PFG-NMR can be found in Chapter 2, Section 2.1.1).

Techniques that involve PFG-NMR can be separated into those that try to extract physical information about the size of molecules in solution, and those that transform this information into a two-dimensional plot in attempt to resolve solution-phase mixtures schematically. In the latter case, the technique has been referred to as diffusion-ordered spectroscopy (DOSY), and is useful in artificially separating mixtures in which molecules interact (Figure 1.13). For example, Frish et al. used DOSY to analyse a competitive host-guest system in which ferrocene and cobaltacene were mixed with calix[4]arenes, and characterise the resultant cobaltacene–calixarene complex.[123] As a result of encapsulation, the diffusion coefficient of the cobaltacene is observed to decrease from 7.3×10^{-10} m^2 s^{-1} to 2.3×10^{-10} m^2 s^{-1}. This diffusion coefficient is identical to that of the host calixarene complex, indicating that the guest is strongly bound to the host, and that exchange between bound and unbound states is slow.[124]

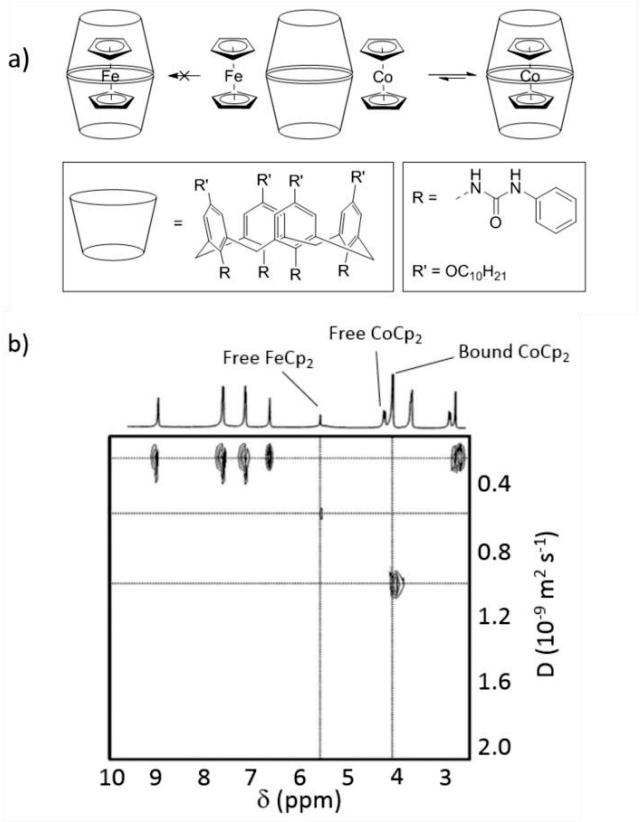


Figure 1.13. (a) Schematic representation of the reaction between a calix[4]arene dimer and two metal complexes, ferrocene and cobaltocene; (b) DOSY spectrum of the crude reaction mixture, adapted from reference 121 copyright © 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

In Chapter 3, Section 3.2.7, we use PFG-NMR to characterise a host-guest complex in a porous liquid, in which the guests are in fast exchange between bound and unbound states. We advance the technique used by Hermans et al.[125] who had used PFG-NMR to calculate proportional occupancy of a host in a fast exchange regime, and we enable this to be carried out between samples of different viscosity.

1.3 Porous polymeric materials

Most porous polymeric materials are insoluble. Like covalent cages, these materials are often synthesised from organic molecules containing light, non-metallic elements such as carbon, hydrogen, nitrogen and oxygen.[126] Many are synthesised by

crosslinking rigid starting materials (Figure 1.14), although the diversity of the underlying materials allows a broad range of synthetic methods to be deployed. Porous aromatic frameworks (PAFs),[127] porous polymer networks (PPNs),[128] conjugated microporous polymers (CMPs)[129,130] and element–organic frameworks (EOFs)[131,132] are shown in Figure 1.14. In addition, Friedel−Crafts reactions have been used to "knit" molecules together, forming polymers with ultra-small pores for H_2 storage,[133] and triazine frameworks have been formed by ionothermal syntheses for molecular separations of CO_2.[134,135]

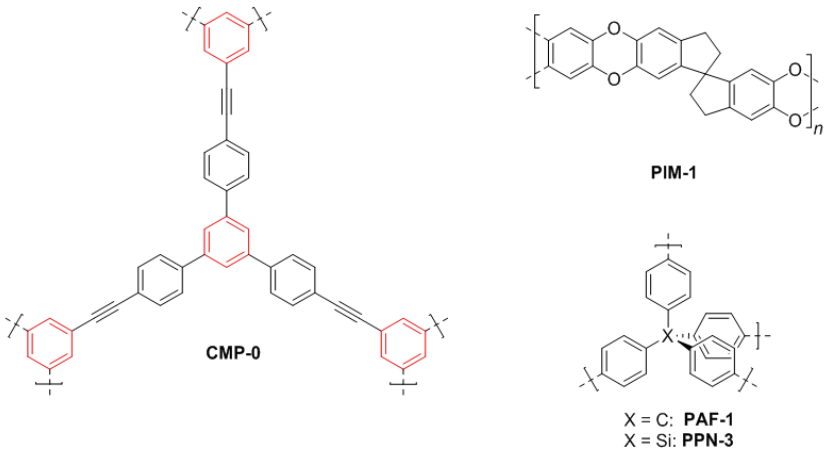


Figure 1.14. Molecular structure of CMP-0, PIM-1, PAF-1 and PPN-3.

Ultrahigh surface areas were generated in the case of **PAF-1**,[127] which had a BET surface area of 5600 $m^2\ g^{-1}$ and **PPN-4** (X = Ge, Figure 1.14), [128] which had a BET surface area of 6461 $m^2\ g^{-1}$. However, these materials are insoluble, which was attributed to their cross-linked structure. CMPs have been reported with more moderate surface areas, between 500 and 1100 $m^2\ g^{-1}$, but have demonstrated utility in catalysis[136,137] and light harvesting.[138,139] CMPs were also found to be insoluble.[129,130] PIMs, in which porosity is generated rigid polymer chains packing

poorly and generating extrinsic space,[16,50] were designed to mimic the pore structures of activated charcoals,[50] and have demonstrated utility as molecular sieves. Particularly, they have excelled as sieving membranes cast from solution for gas separations,[140] and are exceptional amongst porous polymers, because they remain soluble after polymerisation. More generally, solution processability can be advantageous:[141] membrane separation plants can be more energy-efficient,[141] and have a smaller footprint[142] than other types of facilities, such as amine stripping plants for CO_2 capture or gas separation. However, polymer membranes currently incur higher capital costs.[143]

Analysis of polymers in solution is not isolated to post-synthetic characterisation. The rich synthetic diversity surrounding polymer synthesis has enabled preparation of materials that can be synthesised directly into moulds to form monoliths,[144–147] foams[148] and films (Figure 1.15).[141–143,149] Frequently, once cast, the form of the material is tied to its function, so re-processing, such as dissolution for analysis, may alter their properties. However, we reason that an ability to analyse these materials during casting, or under replicated conditions would be beneficial in understanding the polymerisation as it evolves towards the functional material.

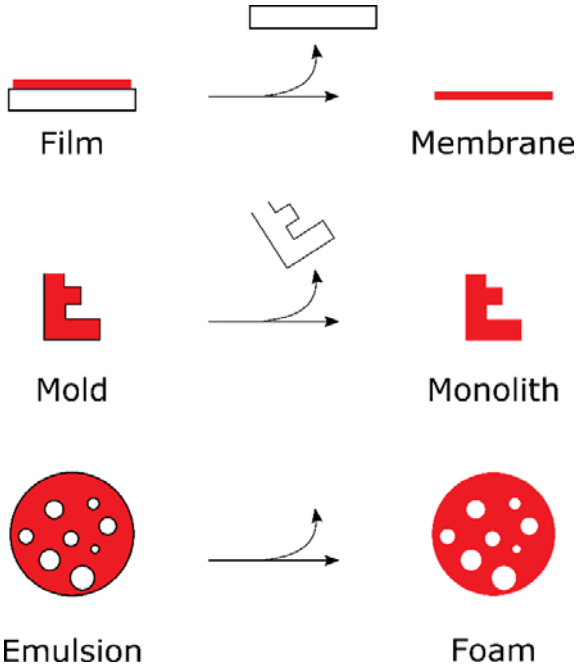


Figure 1.15. Schematic representation of polymer synthesis as membranes (top), monoliths (middle) and foams (bottom).

Molar-mass dispersity ($Đ_M$),[150] a measure of the distribution of molecular weights following polymerisation, is a crucial parameter that can be used to describe the properties of polymers once cast.[151–154] In Chapter 3, we develop a novel method of measuring $Đ_M$ *in situ* using PFG-NMR. This enables properties of polymers to be analysed during synthesis, removing the need for workup and sampling. This method is designed to minimise the mathematical complexity associated with calculating $Đ_M$ and we hope that this will make the process accessible to a wide range of polymer scientists.

1.3.1 Combining the benefits of polymers and molecules

Under certain circumstances, polymeric materials demonstrate several advantages over framework solids, and porous molecular materials. Several porous polymers can

be processed in solution,[16,50] and many demonstrate high thermal stability.[155,156] However, controlling the size and distribution of pores in porous polymeric materials can be challenging.[157,158] Pore heterogeneity has been reported in amorphous polymers, as pores are generated by poor packing rather than molecularly-defined cavities.[158–162] Conversely, porous molecular materials, which are monodisperse, have defined cavity sizes and pore diameters.[163,164] We hypothesised that incorporation of molecules with intrinsic porosity into a processable polymer would combine the benefits of porous molecular materials with those of polymers.

1.4 Conclusions

The majority of the work in this Thesis revolves around developing novel methods for extracting useful information from PFG-NMR that has previously been unreliable or unavailable. In Chapter 2, we describe the implementation of PFG-NMR into a high-throughput screen, and extend the work of Hermans et al.[125] in allowing host-guest chemistry to be compared between several systems with different properties. In Chapter 3, we describe attempts to synthesise intrinsically porous monomers, for the incorporation into a polymeric material. We take significant steps towards controlling the kinetics of post-synthetic functionalisation reactions of intrinsically-porous molecular cages. Using this process, we synthesise mixtures of molecular cages in which the main component is a mono-acylated covalent cage. We suggest that this could form the basis of a 'porous monomer' in future work, which could then be incorporated into a porous polymer. In Chapter 4, we develop a novel method of analysing the PFG-NMR data of polymeric samples, and we enable $Đ_M$ to be measured with unprecedented certainty and ease.

1.5 Bibliography

1 P. A. Wright and P. A. Wright, *Microporous Framework Solids*, Royal Society of Chemistry, Cambridge, 2007.

2 D. Aaron and C. Tsouris, *Sep. Sci. Technol.*, 2005, **40**, 321–348.

3 V. Abetz, T. Brinkmann, M. Dijkstra, K. Ebert, D. Fritsch, K. Ohlrogge, D. Paul, K.-V. Peinemann, S. P. Nunes, N. Scharnagl and M. Schossig, *Adv. Eng. Mater.*, 2006, **8**, 328–358.

4 M. M. Abu-Khader, *Energy Sources, Part A Recover. Util. Environ. Eff.*, 2006, **28**, 1261–1279.

5 P.-P. Knops-Gerrits, D. E. De Vos, E. J. P. Feijen and P. A. Jacobs, *Microporous Mater.*, 1997, **8**, 3–17.

6 J. Lee, J. Li and J. Jagiello, *J. Solid State Chem.*, 2005, **178**, 2527–2532.

7 H. (Hakze) Chon, S.-K. Ihm and Y. S. Uh, *Progress in zeolite and microporous materials : proceedings of the 11th International Zeolite Conference, Seoul, Korea, August 12-17, 1996*, Elsevier Science B.V, 1997.

8 J. I. Paredes, A. Martínez-Alonso and J. M. D. Tascón, *Microporous Mesoporous Mater.*, 2003, **65**, 93–126.

9 S. Coluccia, L. Marchese and G. Martra, *Microporous Mesoporous Mater.*, 1999, **30**, 43–56.

10 F. Thielmann, D. A. Butler, D. R. Williams and E. Baumgarten, *Stud. Surf. Sci. Catal.*, 2000, **129**, 633–638.

11 F. Schüth, K. S. W. Sing and J. Weitkamp, Eds., *Handbook of Porous Solids*, Wiley-VCH, Weinheim, 2008.

12 G. Ferey, *Chem. Soc. Rev.*, 2008, **37**, 191–214.

13 F. Schüth, K. S. W. Sing and J. Weitkamp, *Handbook of Porous Solids*, Wiley-VCH Verlag GmbH, Weinheim, Germany, 2002.

14 A. U. Czaja, N. Trukhan and U. Müller, *Chem. Soc. Rev.*, 2009, **38**, 1284.

15 T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S. Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner and A. I. Cooper, *Nat. Mater.*, 2009, **8**, 973–978.

16 P. M. Budd, B. S. Ghanem, S. Makhseed, N. B. McKeown, K. J. Msayib and C. E. Tattershall, *Chem. Commun.*, 2004, **11**, 230.

17 E. M. Flanigen, *Stud. Surf. Sci. Catal.*, 1991, **58**, 13–34.

18 J. A. Rabo and M. W. Schoonover, *Appl. Catal. A Gen.*, 2001, **222**, 261–275.

19 K. Munakata, T. Fukumatsu, S. Yamatsuki, K. Tanaka and M. Nishikawa, *J. Nucl. Sci. Technol.*, 1999, **36**, 818–829.

20 W. C. Gaul and D. W. Underhill, *Health Phys.*, 2005, **88**, 371–378.

21 J.-R. Li, R. J. Kuppler and H.-C. Zhou, *Chem. Soc. Rev.*, 2009, **38**, 1477.

22 E. D. Bloch, W. L. Queen, R. Krishna, J. M. Zadrozny, C. M. Brown and J. R. Long, *Science*, 2012, **335**, 1606–1610.

23 M. Mastalerz, *Chem. Commun.*, 2008, 4756.

24 A. Goto, K. Sato, Y. Tsujii, T. Fukuda, G. Moad, E. Rizzardo and S. H. Thang, *Macromolecules*, 2001, **34**, 402–408.

25 G. Moad, Y. K. Chong, A. Postma, E. Rizzardo and S. H. Thang, *Polymer*, 2005, **46**, 8458–8468.

26 M. J. Bojdys, S. A. Wohlgemuth, A. Thomas and M. Antonietti, *Macromolecules*, 2010, **43**, 6639–6645.

27 Y.-F. Liang, Y.-R. Zhang and Z.-B. Han, *Zeitschrift fuer Anorg. und Allg. Chemie*, 2012, **638**, 423–426.

28 M. Dincă, A. Dailly, Y. Liu, C. M. Brown, D. A. Neumann and J. R. Long, *J. Am. Chem. Soc.*, 2006, **128**, 16876–16883.

29 N. L. Rosi, *Science*, 2003, **300**, 1127–1129.

30 Y. Li, T. Ben, B. Zhang, Y. Fu and S. Qiu, *Sci. Rep.*, 2013, **3**, 4473–4498.

31 P. A. Wright and J. A. Connor, *Microporous Framework Solids*, Royal Society of Chemistry, Cambridge, 2007.

32 V. Meynen, P. Cool and E. F. Vansant, *Microporous Mesoporous Mater.*, 2007, **104**, 26–38.

33 O. M. Yaghi, M. O'Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi and J. Kim, *Nature*, 2003, **423**, 705–714.

34 A. P. Cote, *Science*, 2005, **310**, 1166–1170.

35 R. Dawson, A. I. Cooper and D. J. Adams, *Prog. Polym. Sci.*, 2012, **37**, 530–563.

36 M. W. Deem and J. M. Newsam, *Nature*, 1989, **342**, 260–262.

37 B. Civalleri, F. Napoli, Y. Noël, C. Roetti and R. Dovesi, *CrystEngComm*, 2006, **8**, 364–371.

38 K. E. Maly, *J. Mater. Chem.*, 2009, **19**, 1781.

39 J. R. Holst, A. Trewin and A. I. Cooper, *Nat. Chem.*, 2010, **2**, 915–920.

40 I. J. Kang, N. A. Khan, E. Haque and S. H. Jhung, *Chem. - A Eur. J.*, 2011, **17**, 6437–6442.

41 T. A. Makal, X. Wang and H.-C. Zhou, *Cryst. Growth Des.*, 2013, **13**, 4760–4768.

42 L. Zhang, K. Chen, B. Chen, J. L. White and D. E. Resasco, *J. Am. Chem. Soc.*, 2015, **137**, 11810–11819.

43 S. LIU, J.-F. Wu, L.-J. Ma, T.-C. Tsai and I. Wang, *J. Catal.*, 1991, **132**, 432–439.

44 P. K. J. Kumar, P. M. Rao and D. L. R. Setty, *Cryst. Res. Technol.*, 1989, **24**, 1027–1033.

45 B. F. Hoskins and R. Robson, *J. Am. Chem. Soc.*, 1990, **112**, 1546–1554.

46 M. P. Suh, H. J. Park, T. K. Prasad and D.-W. Lim, *Chem. Rev.*, 2012, **112**, 782–835.

47 H. Furukawa, N. Ko, Y. B. Go, N. Aratani, S. B. Choi, E. Choi, A. O. O. Yazaydin, R. Q. Snurr, M. O'Keeffe, J. Kim and O. M. Yaghi, *Science*, 2010, **329**, 424–428.

48 X. Feng, X. Ding and D. Jiang, *Chem. Soc. Rev.*, 2012, **41**, 6010–6022.

49 S.-Y. Ding and W. Wang, *Chem. Soc. Rev.*, 2013, **42**, 548–568.

50 N. B. McKeown and P. M. Budd, *Chem. Soc. Rev.*, 2006, **35**, 675–683.

51 S. Kitagawa and R. Matsuda, *Coord. Chem. Rev.*, 2007, **251**, 2490–2509.

52 C. Janiak, *Dalt. Trans.*, 2003, **101**, 2781.

53 S. Kitagawa and K. Uemura, *Chem. Soc. Rev.*, 2005, **34**, 109.

54 M. . Dubinin, *J. Colloid Interface Sci.*, 1967, **23**, 487–499.

55 S. E. Ashbrook, D. M. Dawson and V. R. Seymour, *Phys. Chem. Chem. Phys.*, 2014, **16**, 8223.

56 A. Nossov, F. Guenneau, M.-A. Springuel-Huet, E. Haddad, V. Montouillout, B. Knott, F. Engelke, C. Fernandez and A. Gédéon, *Phys. Chem. Chem. Phys.*, 2003, **5**, 145–4483.

57 T. D. Petersen, G. Balakrishnan and C. L. Weeks, *Dalt. Trans.*, 2015, **44**, 2987–2997.

58 G. Kumari, N. R. Patil, V. S. Bhadram, R. Haldar, S. Bonakala, T. K. Maji and C. Narayana, *J. Raman Spectrosc.*, 2016, **47**, 149–155.

59 M. Mastalerz, *Angew. Chem. Int. Ed.*, 2010, **49**, 5042–5053.

60 N. J. N. Young and B. P. B. Hay, *Chem. Commun.*, 2013, **49**, 1354–1379.

61 M. Mastalerz, *Synlett*, 2013, **24**, 781–786.

62 D. J. Tranchemontagne, Z. Ni, M. O'Keeffe and O. M. Yaghi, *Angew. Chem. Int. Ed.*, 2008, **47**, 5136–5147.

63 M. Mastalerz and I. M. Oppel, *Angew. Chem. Int. Ed.*, 2012, **51**, 5252–5255.

64 S. Jiang, A. Trewin and A. I. Cooper, in *Discovering the Future of Molecular Sciences*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2014, pp. 329–347.

65 G. R. Desiraju, *Nature*, 2001, **412**, 397–400.

66 G. R. Desiraju, *Angew. Chem. Int. Ed.*, 2007, **46**, 8342–8356.

67 J.-M. Lehn, *Chem. Soc. Rev.*, 2007, **36**, 151–160.

68 T. Hasell, J. L. Culshaw, S. Y. Chong, M. Schmidtmann, M. A. Little, K. E. Jelfs, E. O. Pyzer-Knapp, H. Shepherd, D. J. Adams, G. M. Day and A. I. Cooper, *J. Am. Chem. Soc.*, 2014, **136**, 1438–1448.

69 J. T. A. Jones, D. Holden, T. Mitra, T. Hasell, D. J. Adams, K. E. Jelfs, A. Trewin, D. J. Willock, G. M. Day, J. Bacsa, A. Steiner and A. I. Cooper, *Angew. Chem. Int. Ed.*, 2011, **50**, 749–753.

70 M. A. Little, S. Y. Chong, M. Schmidtmann, T. Hasell and A. I. Cooper, *Chem. Commun.*, 2014, **50**, 9465.

71 E. O. Pyzer-Knapp, H. P. G. Thompson, F. Schiffmann, K. E. Jelfs, S. Y. Chong, M. A. Little, A. I. Cooper and G. M. Day, *Chem. Sci.*, 2014, **5**, 2235.

72 T. Hasell, H. Zhang and A. I. Cooper, *Adv. Mater.*, 2012, **24**, 5732–5737.

73 Q. Song, S. Jiang, T. Hasell, M. Liu, S. Sun, A. K. Cheetham, E. Sivaniah and A. I. Cooper, *Adv. Mater.*, 2016, **28**, 2629–2637.

74 M. Brutschy, M. W. Schneider, M. Mastalerz and S. R. Waldvogel, *Adv. Mater.*, 2012, 6049–6052.

75 J. R. Holst and A. I. Cooper, *Adv. Mater.*, 2010, **22**, 5212–5216.

76 M. Mastalerz, *Chem. - A Eur. J.*, 2012, **18**, 10082–10091.

77 B. T. Ibragimov and S. A. Talipov, *J. Incl. Phenom. Mol. Recognit. Chem.*, **17**, 325–328.

78 A. I. Cooper, *Angew. Chem. Int. Ed.*, 2012, **51**, 7892–7894.

79 G. R. Desiraju, *Crystal engineering: the design of organic solids*, Elsevier Science Ltd., 1st edn., 1989.

80 B. T. Ibragimov, S. a. Talipov and T. F. Aripov, *J. Incl. Phenom. Mol. Recognit. Chem.*, 1994, **17**, 317–324.

81 R. D. Chapman, R. D. Gilardi, M. F. Welker and C. B. Kreutzberger, *J. Org. Chem.*, 1999, **64**, 960–965.

82 F. Imashiro, M. Yoshimura and T. Fujiwara, *Acta Crystallogr., Sect. C Cryst. Struct. Commun.*,

1998, **54**, 1357–1360.

83 P. Sozzani, A. Comotti, R. Simonutti, T. Meersmann, J. W. Logan and A. Pines, *Angew. Chem. Int. Ed.*, 2000, **39**, 2695–2699.

84 P. Sozzani, S. Bracco, A. Comotti, L. Ferretti and R. Simonutti, *Angew. Chem. Int. Ed.*, 2005, **44**, 1816–1820.

85 G. Couderc, T. Hertzsch, N.-R. Behrnd, K. Krämer and J. Hulliger, *Microporous Mesoporous Mater.*, 2006, **88**, 170–175.

86 A. I. Cooper, *Angew. Chem. Int. Ed.*, 2011, **50**, 996–998.

87 G. Zhang, O. Presly, F. White, I. M. Oppel and M. Mastalerz, *Angew. Chem. Int. Ed.*, 2014, **53**, 1516–1520.

88 Y. Jin, Q. Wang, P. Taynton and W. Zhang, *Acc. Chem. Res.*, 2014, **47**, 1575–1586.

89 P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Dynamic Combinatorial Chemistry*, Wiley-VCH, Weinheim, 1st edn., 2010, vol. 106.

90 G. Gasparini, M. Dal Molin, A. Lovato and L. J. Prins, in *Supramolecular Chemistry*, John Wiley & Sons, Ltd, 2012.

91 J.-M. Lehn, *Science*, 2001, **291**, 2331–2332.

92 Y. Jin, C. Yu, R. J. Denman and W. Zhang, *Chem. Soc. Rev.*, 2013.

93 S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders and J. F. Stoddart, *Dynamic covalent chemistry.*, 2002, vol. 41.

94 A. Herrmann, *Chem. Soc. Rev.*, 2014, **43**, 1899–933.

95 C. D. Meyer, C. S. Joiner and J. F. Stoddart, *Chem. Soc. Rev.*, 2007, **36**, 1705.

96 Y. Liu and Z.-T. Li, *Aust. J. Chem.*, 2013, **66**, 9–22.

97 Y. Jin, Y. Zhu and W. Zhang, *CrystEngComm*, 2013, **15**, 1484.

98 A. Khalaji, S. Maghsodlou Rad, G. Grivani, K. Fejfarova, M. Dusek and D. Das, *J. Chem. Crystallogr.*, 2011, **41**, 1145–1149.

99 T. Mitra, X. Wu, R. Clowes, J. T. A. Jones, K. E. Jelfs, D. J. Adams, A. Trewin, J. Bacsa, A. Steiner and A. I. Cooper, *Chem. - A Eur. J.*, 2011, **17**, 10235–10240.

100 D. P. Lydon, N. L. Campbell, D. J. Adams and A. I. Cooper, *Synth. Commun.*, 2011, **41**, 2146–2151.

101 J. T. A. Jones, T. Hasell, X. Wu, J. Bacsa, K. E. Jelfs, M. Schmidtmann, S. Y. Chong, D. J.

Adams, A. Trewin, F. Schiffman, F. Cora, B. Slater, A. Steiner, G. M. Day and A. I. Cooper, *Nature*, 2011, **474**, 367–371.

102 M. W. Schneider, I. M. Oppel, H. Ott, L. G. Lechner, H.-J. S. Hauswald, R. Stoll and M. Mastalerz, *Chem. - A Eur. J.*, 2012, **18**, 836–847.

103 K. E. Jelfs, X. Wu, M. Schmidtmann, J. T. A. Jones, J. E. Warren, D. J. Adams and A. I. Cooper, *Angew. Chem. Int. Ed.*, 2011, **50**, 10653–10656.

104 K. E. Jelfs, E. G. B. Eden, J. L. Culshaw, S. Shakespeare, E. O. Pyzer-Knapp, H. P. G. Thompson, J. Bacsa, G. M. Day, D. J. Adams, A. I. Cooper, O. Edward, H. P. G. Thompson, J. Bacsa, G. M. Day, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2013, **135**, 9307–9310.

105 V. Santolini, G. A. Tribello and K. E. Jelfs, *Chem. Commun.*, 2015, **51**, 15542–15545.

106 S. Jiang, J. Bacsa, X. Wu, J. T. A. Jones, R. Dawson, A. Trewin, D. J. Adams and A. I. Cooper, *Chem. Commun.*, 2011, **47**, 8919.

107 S. Jiang, J. T. A. Jones, T. Hasell, C. E. Blythe, D. J. Adams, A. Trewin and A. I. Cooper, *Nat. Commun.*, 2011, **2**, 207.

108 G. Zhang and M. Mastalerz, *Chem. Soc. Rev.*, 2014, **43**, 1934–1947.

109 D. Braga, F. Grepioni, E. Tedesco, K. Biradha and G. R. Desiraju, *Organometallics*, 1997, **16**, 1846–1856.

110 A. Nangia and G. R. Desiraju, *Chem. Commun.*, 1999, 605–606.

111 E. R. T. Tiekink and J. Zukerman-Schpector, Eds., *The Importance of Pi-Interactions in Crystal Engineering*, John Wiley & Sons, Ltd, Chichester, UK, 2012.

112 J. A. R. P. Sarma and G. R. Desiraju, *Acc. Chem. Res.*, 1986, **19**, 222–228.

113 J. D. Dunitz and A. Gavezzotti, *Angew. Chem. Int. Ed.*, 2005, **44**, 1766–1787.

114 G. R. Desiraju, *J. Mol. Struct.*, 2003, **656**, 5–15.

115 D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, *Acta Crystallogr. Sect. B Struct. Sci.*, 2011,

67, 535–551.

116 G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallogr. Sect. B Struct. Sci.*, 2009, **65**, 107–125.

117 G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta Crystallogr. Sect. B Struct. Sci.*, 2005, **61**, 511–527.

118 W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr. Sect. B Struct. Sci.*, 2002, **58**, 647–661.

119 J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallogr. Sect. B Struct. Sci.*, 2000, **56**, 697–714.

120 C. S. Johnson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1999, **34**, 203–256.

121 L. Avram and Y. Cohen, *Chem. Soc. Rev.*, 2015, **44**, 586–602.

122 A. Einstein, *Ann. Phys.*, 1905, **322**, 549–560.

123 L. Frish, M. O. Vysotsky, V. Böhmer and Y. Cohen, *Org. Biomol. Chem.*, 2003, **1**, 2011–2014.

124 A. Gafni and Y. Cohen, *J. Org. Chem.*, 1997, **62**, 120–125.

125 T. M. Hermans, M. A. C. Broeren, N. Gomopoulos, P. van der Schoot, M. H. P. van Genderen, N. A. J. M. Sommerdijk, G. Fytas and E. W. Meijer, *Nat. Nanotechnol.*, 2009, **4**, 721–726.

126 J. Germain, J. M. J. Fréchet and F. Svec, *Small*, 2009, **5**, 1098–1111.

127 T. Ben, H. Ren, S. Ma, D. Cao, J. Lan, X. Jing, W. Wang, J. Xu, F. Deng, J. M. Simmons, S.

Qiu and G. Zhu, *Angew. Chem. Int. Ed.*, 2009, **48**, 9457–9460.

128 D. Yuan, W. Lu, D. Zhao and H.-C. Zhou, *Adv. Mater.*, 2011, **23**, 3723–3725.

129 J.-X. Jiang, F. Su, A. Trewin, C. D. Wood, N. L. Campbell, H. Niu, C. Dickinson, A. Y. Ganin, M. J. Rosseinsky, Y. Z. Khimyak and A. I. Cooper, *Angew. Chem. Int. Ed.*, 2007, **46**, 8574–8578.

130 J.-X. Jiang, F. Su, A. Trewin, C. D. Wood, H. Niu, J. T. A. Jones, Y. Z. Khimyak and A. I. Cooper, *J. Am. Chem. Soc.*, 2008, **130**, 7710–7720.

131 M. Rose, N. Klein, W. Böhlmann, B. Böhringer, S. Fichtner and S. Kaskel, *Soft Matter*, 2010, **6**, 3918.

132 M. Rose, W. Böhlmann, M. Sabo and S. Kaskel, *Chem. Commun.*, 2008, **43**, 2462.

133 J. Germain, J. Hradil, J. M. J. Fréchet and F. Svec, *Chem. Mater.*, 2006, **18**, 4430–4435.

134 P. Kuhn, M. Antonietti and A. Thomas, *Angew. Chem. Int. Ed.*, 2008, **47**, 3450–3453.

135 X. Zhu, C. Tian, S. M. Mahurin, S.-H. Chai, C. Wang, S. Brown, G. M. Veith, H. Luo, H. Liu and S. Dai, *J. Am. Chem. Soc.*, 2012, **134**, 10478–10484.

136 J.-X. Jiang, C. Wang, A. Laybourn, T. Hasell, R. Clowes, Y. Z. Khimyak, J. Xiao, S. J. Higgins, D. J. Adams and A. I. Cooper, *Angew. Chem. Int. Ed.*, 2011, **50**, 1072–1075.

137 L. Chen, Y. Yang and D. Jiang, *J. Am. Chem. Soc.*, 2010, **132**, 9138–9143.

138 J.-X. Jiang, A. Trewin, D. J. Adams and A. I. Cooper, *Chem. Sci.*, 2011, **2**, 1777.

139 L. Chen, Y. Honsho, S. Seki and D. Jiang, *J. Am. Chem. Soc.*, 2010, **132**, 6742–6748.

140 P. M. Budd, E. S. Elabas, B. S. Ghanem, S. Makhseed, N. B. McKeown, K. J. Msayib, C. E. Tattershall and D. Wang, *Adv. Mater.*, 2004, **16**, 456–459.

141 D. F. Sanders, Z. P. Smith, R. Guo, L. M. Robeson, J. E. McGrath, D. R. Paul and B. D. Freeman, *Polymer*, 2013, **54**, 4729–4761.

142 R. W. Baker and K. Lokhandwala, *Ind. Eng. Chem. Res.*, 2008, **47**, 2109–2121.

143 R. W. Baker, *Membrane Technology and Applications*, Wiley, Chichester, UK, Third., 2012.

144 B. Beiler, Á. Vincze, F. Svec and Á. Sáfrány, *Polymer*, 2007, **48**, 3033–3040.

145 M. Rose, N. Klein, I. Senkovska, C. Schrage, P. Wollmann, W. Böhlmann, B. Böhringer, S. Fichtner and S. Kaskel, *J. Mater. Chem.*, 2011, **21**, 711–716.

146 F. Svec, *J. Chromatogr. A*, 2010, **1217**, 902–924.

147 C. D. Wood, B. Tan, A. Trewin, F. Su, M. J. Rosseinsky, D. Bradshaw, Y. Sun, L. Zhou and

A. I. Cooper, *Adv. Mater.*, 2008, **20**, 1916–1921.

148 P. S. Liu, G. F. Chen, P. S. Liu and G. F. Chen, in *Porous Materials*, 2014, pp. 345–382.

149 R. W. Baker, *Ind. Eng. Chem. Res.*, 2002, **41**, 1393–1411.

150 R. F. T. Stepto, *Pure Appl. Chem.*, 2009, **81**, 351–353.

151 D. Broseta, G. H. Fredrickson, E. Helfand and L. Leibler, *Macromolecules*, 1990, **23**, 132–139.

152 M. J. Struglinski and W. W. Graessley, *Macromolecules*, 1985, **18**, 2630–2643.

153 D. M. Cooke and A.-C. Shi, *Macromolecules*, 2006, **39**, 6661–6671.

154 L. Lu, T. Zheng, T. Xu, D. Zhao and L. Yu, *Chem. Mater.*, 2015, **27**, 537–543.

155 L.-B. Sun, A.-G. Li, X.-D. Liu, X.-Q. Liu, D. Feng, W. Lu, D. Yuan and H.-C. Zhou, *J. Mater. Chem. A*, 2015, **3**, 3252–3256.

156 X. Wu, R. Wang, H. Yang, W. Wang, W. Cai and Q. Li, *J. Mater. Chem. A*, 2015, **3**, 10724–10729.

157 C. Weder, *Angew. Chem. Int. Ed.*, 2008, **47**, 448–450.

158 K. Matyjaszewski, *Science*, 2011, **333**, 1104–1105.

159 N. B. McKeown and P. M. Budd, *Macromolecules*, 2010, **43**, 5163–5176.

160 S. A. Al-Muhtaseb and J. A. Ritter, *Adv. Mater.*, 2003, **15**, 101–114.

161 M. P. Tsyurupa and V. A. Davankov, *React. Funct. Polym.*, 2006, **66**, 768–779.

162 M. Mastalerz, *Angew. Chem. Int. Ed.*, 2008, **47**, 445–447.

163 D. Holden, K. E. Jelfs, A. I. Cooper, A. Trewin and D. J. Willock, *J. Phys. Chem. C*, 2012, **116**, 16639–16651.

164 L. Chen, P. S. Reiss, S. Y. Chong, D. Holden, K. E. Jelfs, T. Hasell, M. A. Little, A. Kewley, M. E. Briggs, A. Stephenson, K. M. Thomas, J. A. Armstrong, J. Bell, J. Busto, R. Noel, J. Liu, D. M. Strachan, P. K. Thallapally and A. I. Cooper, *Nat. Mater.*, 2014, **13**, 954–60.

Chapter 2. Structural information from Diffusion NMR

2.1 Introduction

This Chapter surrounds the use of pulsed field gradient (PFG) NMR techniques, and their use in measuring the diffusion speed of molecules in solution. This is accomplished by increasing the applied magnetic field across the length of a sample, which allows molecules to be encoded location-specifically.

2.1.1 Diffusion NMR - theory

Inside an NMR magnet, the sample is surrounded by a coil that generates a small magnetic field, *g*, along the same axis as the permanent magnet.[1] For simplicity, this is considered the *z*-axis. NMR-active nuclei, such as the proton in hydrogen, align with the magnetic field. They precess about their axis at a rate, ω, that is proportional to magnetic field strength felt by the nucleus, B_{eff} (Equation 1)[2]:

$$\omega = B_{eff}\gamma \qquad (1)$$

where γ is the gyromagnetic ratio of the nucleus and $B_{eff} = B_0 + g$, where B_0 is the magnitude of the static magnetic field and g is the amplitude of the magnetic field created by the *z*-gradient coil. The magnitude of g depends on the position along the z-axis. As g is proportional to z, B_{eff} increases across the sample, and nuclei in one position will spin at a different rate to those at another. This causes dephasing, which can be thought of as interrupting the alignment of the NMR spins in such a way that generates destructive interference. This is observed as loss of NMR signal (Figure 2.1).[2]

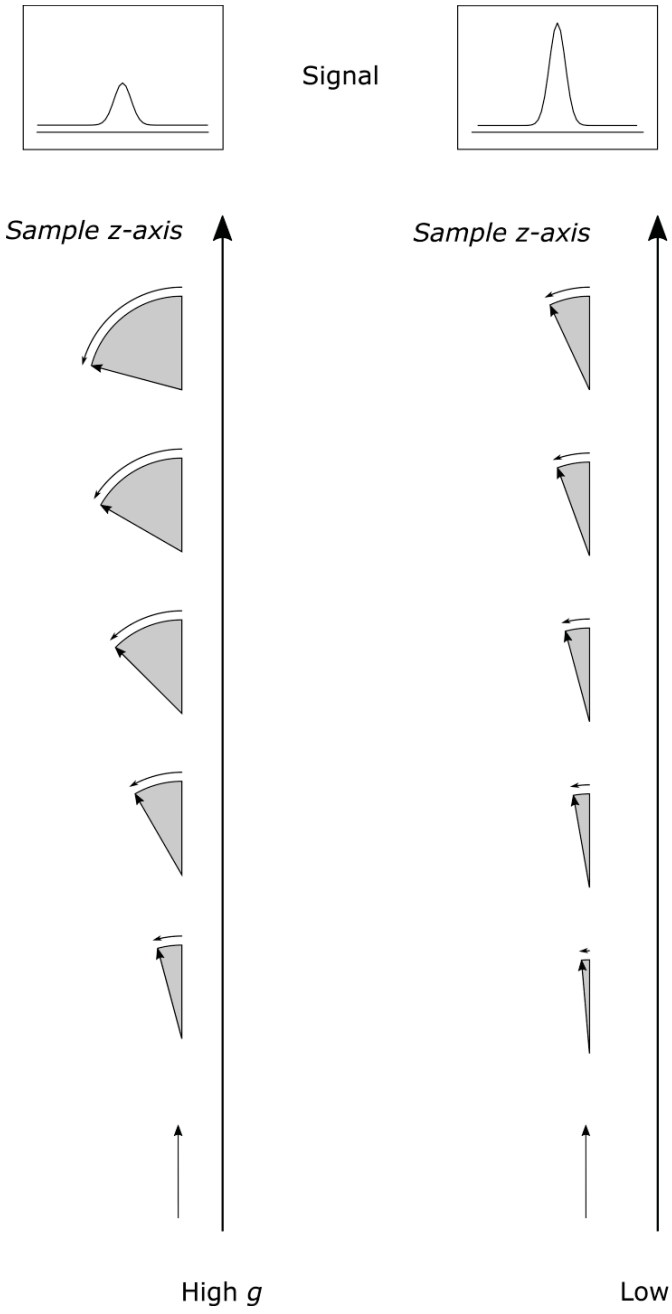


Figure 2.1. A schematic depiction of signal de-phasing with increasing gradient amplitude. Examples of this are shown for both high (*top*) and low (*bottom*) gradient amplitudes.

This dephasing gradient is applied once at the beginning of the NMR experiment, and once after a period of time has passed to allow particles to diffuse within the solution. This time period is referred to as the diffusion encoding delay, Δ, and allows molecules to move along the z-axis between pulses. The initial application of the z-

gradient pulse (g_1) encodes NMR-active nucleii with a location-specific dephasing (Figure 2.1)

In a standard Stimulated-Echo (STE) experiment, radio-frequency pulses are also applied during this time, which causes phase-differences accumulated during the first gradient pulse to interact with the phase-differences accumulated during the second (Figure 2.2a). If a particle has not moved at all during Δ, the second application of the z-gradient pulse (g_2) will decrease dephasing by the same quantity as it has been increased. In cases where no particles in the sample have diffused, dephasing is cancelled, and 100 % of the signal is recovered.

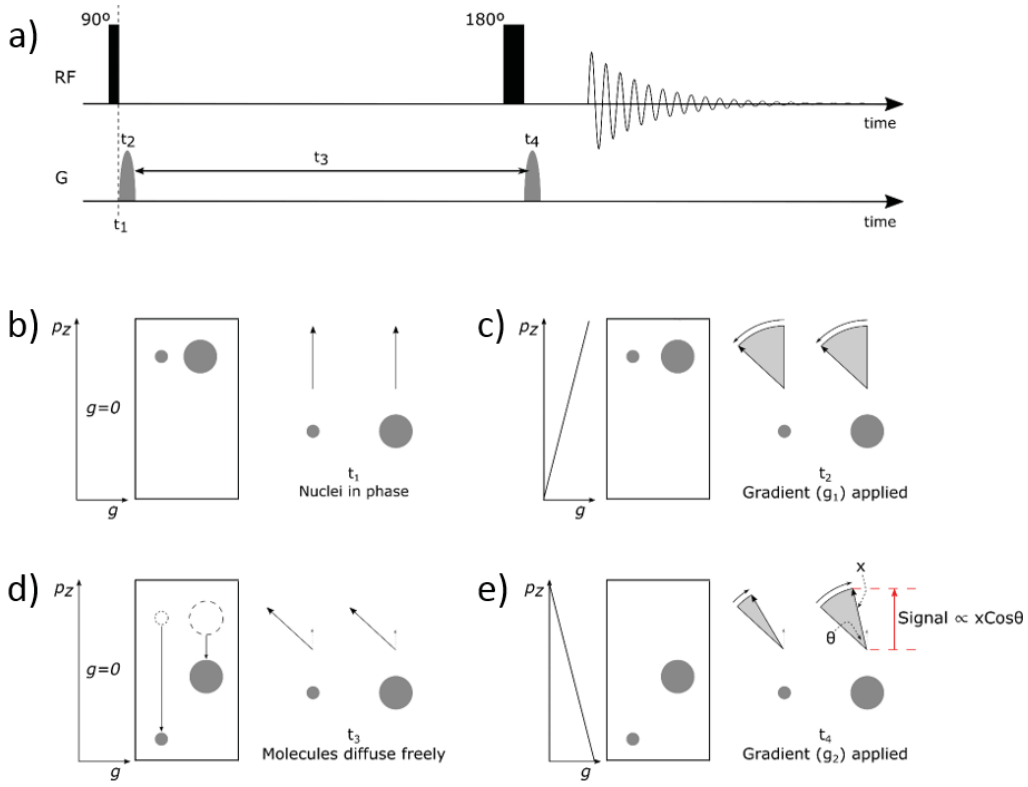


Figure 2.2. (a) A schematic representation of an STE PFG-NMR experiment. Timelines for RF and gradient (G) pulses are shown. RF pulses are labelled with pulse lengths; (b – e) z-gradient amplitude is shown as a graph (*left*), and is only active during t_2 and t_4. A schematic representation of two particles is shown (*centre*). A schematic example of signal phase is shown as an arrow, which is rotated by a magnetic field gradient. The proportion of signal measured during acquisition is shown in red.

Because the quantity of phase change is determined by B_{eff}, particles that have moved from an area of high B_{eff} to an area of low B_{eff} during Δ will be partially re-phased. Particles that have moved a great distance generate a smaller NMR signal than those that have moved more slowly (Figure 2.3).[3,4] Diffusion path lengths from nanometres to micrometres in scale can be measured,[5–7] which facilitates characterisation of dissolved species that range from angstroms to tens of nanometres in size.[8,9]

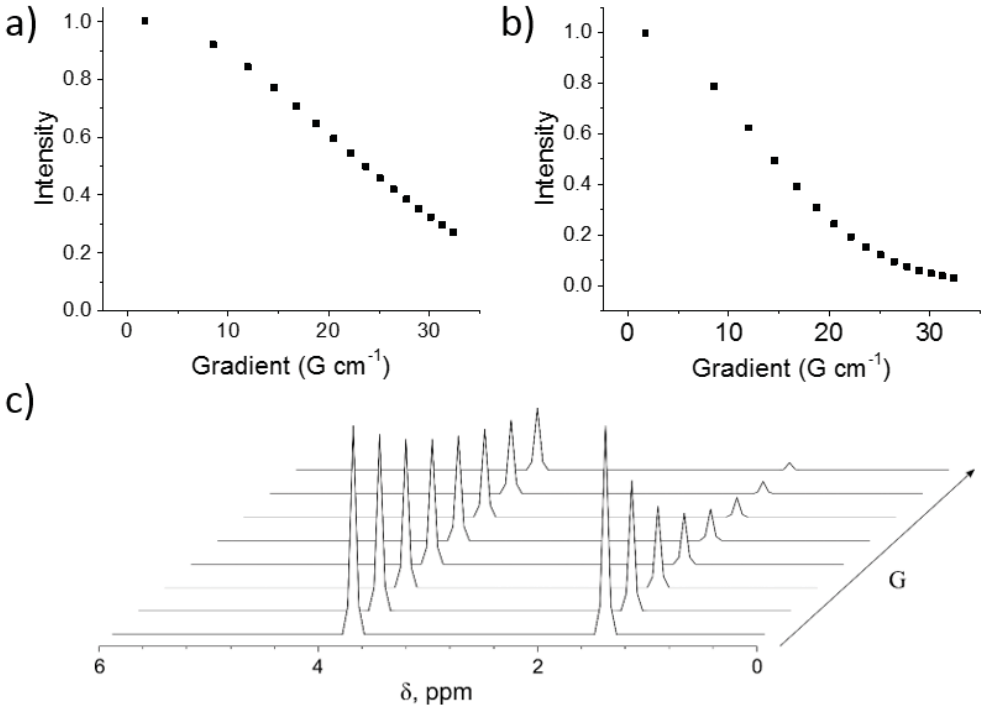


Figure 2.3. The signal intensity of (a) slow- and (b) fast-moving particles in solution with increasing gradient intensity; (c) A schematic representation of a PFG-NMR experiment, focussed around the up-field area of a ^{1}H NMR spectrum. The left hand peak corresponds to the slow-moving particle in (a). The right-hand peak corresponds to the fast-moving particle in (b).

As molecules are not uniformly orientated, and tumble through solution randomly (except under certain specific conditions),[10,11] the measurements result from the *length* of the diffusion path, and not the manner of travel. Furthermore, information about size must be derived using certain approximations, because data about molecular motion is orientationally-averaged. The most basic of these assumptions is that molecules are spherical,[12,13] which enables us to assume that they diffuse with equal speed in all directions. The resultant velocity vector has a magnitude, but no direction, and this magnitude, D, is inversely proportional to the solvodynamic radius, r_s (Equation 2).

$$r_S \propto \frac{1}{D} \qquad (2)$$

This diffusion coefficient, which describes the speed of diffusion, is related to the solvodynamic radius, r_s using the Stokes-Einstein equation (Equation 3).[12,13]

$$r_S = \frac{kT}{6\pi\eta D} \quad (3)$$

where k is the Boltzmann constant, T is the absolute temperature, and η is the viscosity of the solvent. Solute particles are assumed to be non-interacting, which is discussed in Section 2.1.2.

Several factors are important in ensuring accurate data is measured.[7,14,15] In cases where the viscosity of the solvent is changed by the solute, accurate measurements of viscosity are required to calculate molecular size (Equation 3). Furthermore, where two systems of different viscosity are measured, diffusion coefficients can only be compared when normalised by this change.

Small, charged species have been reported to diffuse with a permanent hydration sphere, which can make particles appear greater in size.[16,17] However, in the case of large, neutral species, which interact with solvent *via* Van der Waals (VdW) forces, solvodynamic radii are more consistent with molecular dimensions.[18] In this dilute, neutral regime, molecular dimensions are calculated assuming that no intermolecular interactions affect the speed of diffusion through solution.

The concentration of solute, which can also affect the viscosity, is an important factor, because the Stokes-Einstein equation (Equation 3) assumes solute molecules do not interact. At high concentrations, where this may not be true, diffusion coefficients can become misleading (see Section 2.1.2).[15] Furthermore, molecular anisotropy is important in understanding the meaning of data obtained by PFG-NMR, which is discussed in detail below (see Section 2.1.4).

2.1.2 Concentration

The mean free path of a molecule describes the distance a molecule travels before its trajectory is altered by collision with another molecule. Collision frequency is proportional to the density of particles (*n*) as well as the cross-sectional collision area (σ), a factor that depends on the radii of both colliding objects.[19] The distance between collisions, *l*, is then (Equation 4):

$$l = (n \cdot \sigma_{avg})^{-1} \qquad (4)$$

where σ_{avg} is the average cross-sectional collision area governing diffusion of the particle under study. The cross-sectional area of collisions is determined by the sum of the radii of the colliding particles. Figure 2.4 depicts two examples of collisions that occur in a solution that contains a macromolecular solute. In this example, the macromolecule is larger than the solvent molecule, which is often the case.[18,20] The collisional cross-section of solvent-solute interactions ($\sigma_{\mathbf{v\text{-}u}}$, where **v** indicates solvent and **u** indicates solute) is smaller than that of solute-solute interactions, $\sigma_{u\text{-}u}$ (Figure 2.4).

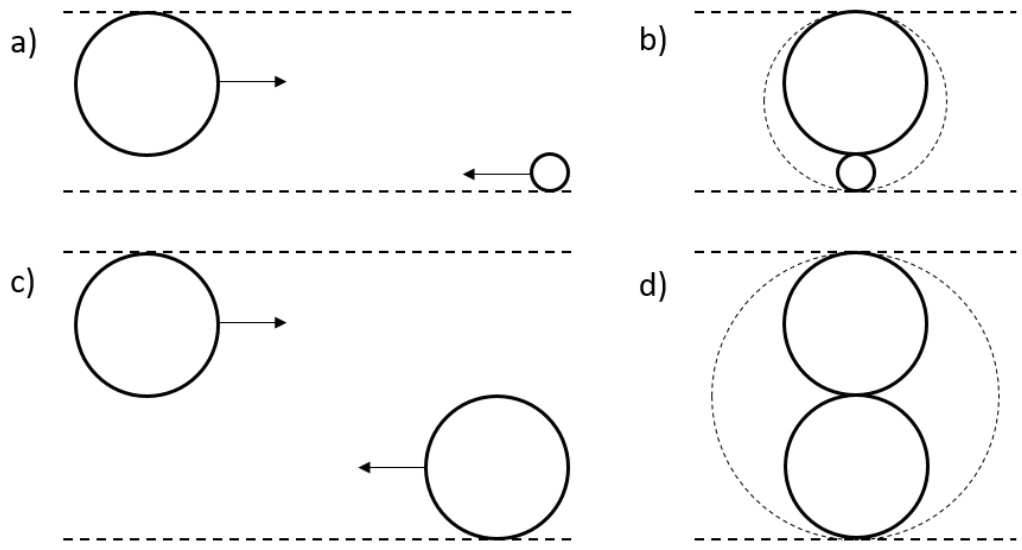


Figure 2.4. Collision paths (a and c) and cross-sectional areas (b and d) for combinations of solute and solvent particles. Cross-sectional collision areas are represented by dotted black circles. Solvent and solute molecules are idealised as small and large circles respectively.

The average collisional cross-section, is calculated as the average over all collisions. Thus, at infinite dilution, as solute particles are surrounded by an infinite number of solvent particles, $\sigma_{avg} \approx \sigma_{u\text{-}v}$. It is assumed here, for simplicity, that cross-sectional areas are independent of molecular velocity, although a more thorough consideration of the "collision integral" is possible.[21,22]

As the concentration of solute particles increases, the number of solute-solute interactions increases and σ_{avg} approaches $\sigma_{u\text{-}u}$. A particle that collides, and is re-directed by its collision, will have a smaller overall displacement than one which is unimpeded.[23] The diffusion coefficient, which measures displacement, is thus observed to decrease as the number of collisions increases.[19] The observed diffusion coefficient, D_{obs}, is found to decrease with concentration (Equation 5):

$$D_{obs} = D_0(1 - k\phi) \qquad (5)$$

where D_0 is the diffusion coefficient at infinite dilution, ϕ is the proportional volume taken up by the measured species and k is a correction factor.[23,24] The magnitude of this correction factor is dependent on solute-solute interactions, and has been reported between 1.75 and 2.[14,15] The correction factor can be calculated by measuring D_{obs} at known concentrations and deriving $k\phi$ from the rate at which D_{obs} decreases (Figure 2.5). Calculation of ϕ is achieved by calculating the intercept, which corresponds to D_0, and estimating the value of r_s using the Stokes-Einstein equation (Equation 3).

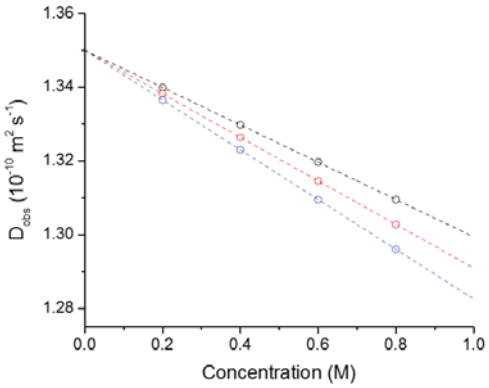


Figure 2.5. The observed diffusion coefficients for three particles of equal size, but differing levels of inter-molecular interaction. D_{obs} is calculated from Equation 5 for a particle where $D_0 = 1.35 \times 10^{-10}$ m^2 s^{-2} using three values of *k*: 1.50 (black), 1.75 (red), 2.00 (blue).

This decrease in D_{obs} can impact experimental decisions, where trade-offs are made between a high signal-to-noise (s/n), observed at high concentrations, and the precision of results, which highest at high dilution. Furthermore, ϕ must be calculated at low enough concentrations that the viscosity of the solution is unaffected by the presence of solute (see Section 2.1.3). Where this is not the case, compensation must be made for viscosity, and the resultant change in D_{obs} fitted to a curve.

2.1.3 Viscosity and convection

Viscosity is an important parameter, because solutions may convect at low viscosity, which can have a significant impact on the diffusion coefficient measured.[25] In the presence of convection, molecular tumbling is often screened by macroscopic movement of the solution on a greater-than-micrometre scale (Figure 2.6).

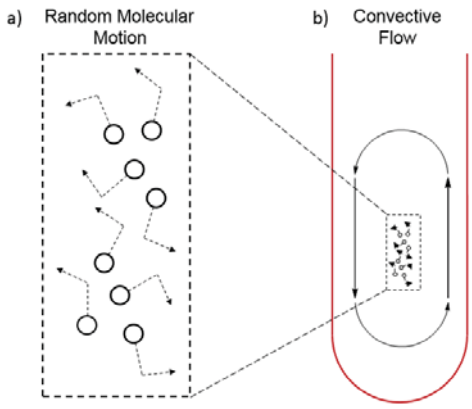


Figure 2.6. Diffusion of particles in the (a) absence and (b) presence of convection. Here, convection is depicted as being on a greater scale than random molecular motion, although this is not always the case.[25]

Convection, the concerted movement of particles in solution, can be compensated for by performing two identical PFG-NMR experiments back-to-back, inverting the phase of the second.[25,26] Diffusion is a random process, and data that results from diffusion is not affected by this treatment. However, convection that adds apparent size to molecules in the first experiment will decrease their apparent size in the second. The average of these effects produces a value that is independent of convection.[25]

Significant problems can also arise as conversion from diffusion coefficients to solvodynamic radii requires η, and increases in solute concentration are known to affect viscosity.[27,28] As with changes of D_{obs} at increased concentration, consideration of viscosity can inform experiment design, because higher concentrations are desirable for a greater signal-to-noise ratio, but can skew experimental results. Thus, an estimate of the rate at which concentration affects D_{obs} is required alongside diffusion data to validate results.

2.1.4 Molecular anisotropy

Where the molecule being measured is anisotropic, that is, it is longer in one dimension than in another, this can affect the magnitude of D.[16] Large molecular structures, which aggregate in a single dimension, for example, appear volumetrically larger than their true size. Figure 2.7 demonstrates this by comparing hypothetical one-dimensional aggregates with their apparent spherical size.

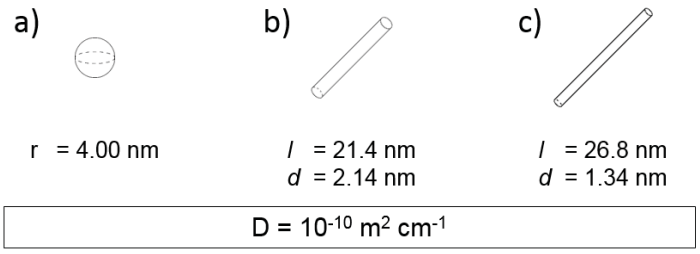


Figure 2.7. Three objects with identical diffusion coefficients; (a) A sphere of radius 4 nm; (b and c) Cylinders with aspect ratios (l/d) of 10 and 20 respectively. The diffusion coefficient of (a) was calculated using Equation 3, diffusion coefficients of (b) and (c) were calculated using Equation 7, below.

Modifications to the Stokes-Einstein equation (Equation 3) that attempt to factor in anisotropy must average the effects of differing diffusion speeds in several directions. A cuboid, for example, of dimensions 1×1×3 will diffuse three times more quickly along its long axis than along its other two (Figure 2.8).[29,30]

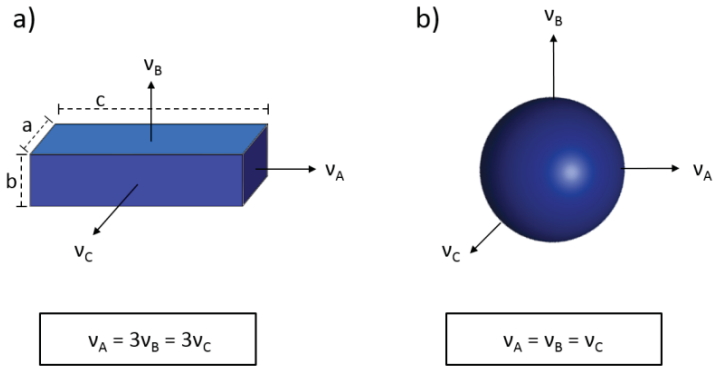


Figure 2.8. Diffusion speeds (*v*) of two solids along the a, b and c axes; (a) a cuboid with dimensions: c = 3a = 3b; (b) a sphere.

Models are often chosen with an emphasis on symmetry; for example, spheroids rather than cuboids were chosen to model amide aggregates by Stjerndahl et al.[16] Under conditions of random molecular motion, axial symmetry (C∞) was found to represent an average image of an anisotropic molecule in solution.[31–33]

In each case, anisotropy results in a dramatic increase in the apparent size of the molecule. Equation 6 demonstrates a common strategy, in which a correction factor α (Equation 6) is introduced, the size of which is determined by the number of dimensions in which diffusion is limited, and the degree to which the molecule deviates from spherical symmetry.[19]

$$D_{obs} = \alpha D \qquad (6)$$

where, here, D may be thought of as the diffusion coefficient in the absence of asymmetrical factors affecting diffusion. Equations for cylindrical (Equation 7)[17] as well as prolate (Equation 8) and oblate (Equation 9) spheroids[16] have been used to characterise macromolecular aggregates in solution by both NMR[34,35] and light-scattering.[36]

$$D_{obs} = \frac{kT}{3\pi\eta l}\,(lnp + \nu) \qquad (7)$$

$$D_{obs} = \frac{kT}{3\pi\eta b}\,\frac{ln\left(r+\sqrt{r^2-1}\right)}{\sqrt{r^2-1}} \qquad (8)$$

$$D_{obs} = \frac{kT}{3\pi\eta b} \frac{arctan(r^2-1)^{1/2}}{(r^2-1)^{1/2}} \quad (9)$$

where *l* is the length of the cylinder, *p* is the ratio between the length and diameter, *b* is the long axis of the ellipsoid, and *r* is the ratio between the long and short axes (these are defined schematically in Figure 2.9a). Sections of each equation that correlate with the scaling factor α have been highlighted in blue.

Cylindrical and spheroidal particles differ in the rate at which the anisotropic correction factor increases (Figure 2.9b–d). Limitations of this approach are observed when the deviation from spherical symmetry is marginal. For example, short cylindrical 'disks' begin to tumble about the short axis when the ratio (r) between the length (*l*) and the diameter (d) of the cylinder is less than 3.[17]

Because anisotropic information is lost during molecular tumbling, it is difficult to determine the shape of a particle from isolated ^{1}H NMR experiments. For this reason, it is important to predict the shape of the molecule or aggregate, and then to fit the proposed dimensions into the appropriate equation. This is explored in Section 2.2.1.4, where computational models of anisotropic cage molecules are compared with PFG-NMR results. In Section 2.2.3., this technique is employed within a high-throughput screen, enabling a diverse range of cages to be characterised *in situ* without the need for workup. Finally, Section 2.2.4 describes how PFG-NMR can be used to characterise host-guest interactions within cage solutions, where cages are able to bind gaseous guest molecules and release them on application of chemical and physical stimuli.

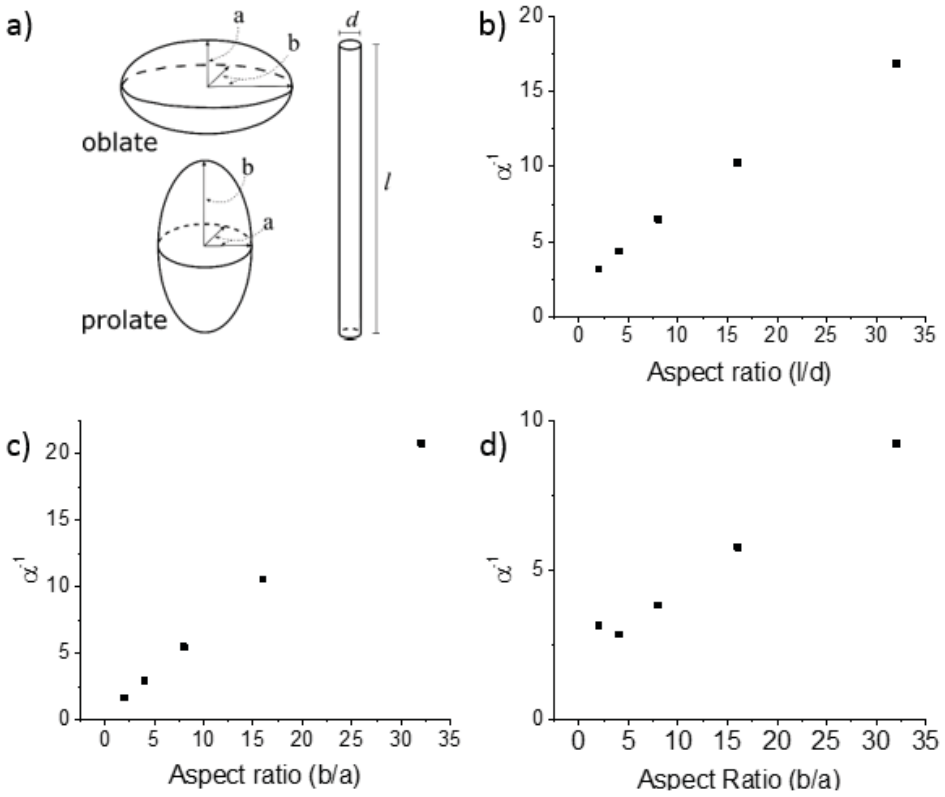


Figure 2.9. (a) The dimensions of cylinders and spheroids, which are used to calculate the aspect ratio of the shape. The reciprocal correction factors for (b) cylinders, (c) oblate spheroids and (d) prolate spheroids are calculated from Equations 7–9.

Molecular cages can have an external surface that is concave rather than convex (see Chapter 1, Section 1.2.3). This is difficult to account for when converting between D_{obs} and r_s, and is compounded when the topology of the cage is unknown. So far, no models have been produced to account for this, and little work has been done to test what real-life property is being measured. Section 2.2.1 deals with two hypotheses, which aim to extract this information from experimental results. Section 2.2.1.4 deals with structures that had been predicted to adopt a prismatic geometry, and an assessment is made as to which spheroidal model fits most appropriately before diffusion NMR data is compared with the computationally-generated structures.

2.1.5 PFG-NMR Analysis of molecules with high symmetry

Some supramolecules, such as tetrahedral organic cages (symmetry T_d), and large icosahedral metallocages (symmetry I_h), have a high degree of symmetry, which is seen to justify their treatment as a solvodynamic sphere.[18] In these cases, solvodynamic radii that lie between 50 and 110 % of the total size of the molecule are reported to be in agreement with experimental results (Figure 2.10).[37–39] To our knowledge, little work has been done to quantify the degree to which molecular shape affects the size of the diffusion coefficient obtained, and how this might be generalised across a number of cage topologies.

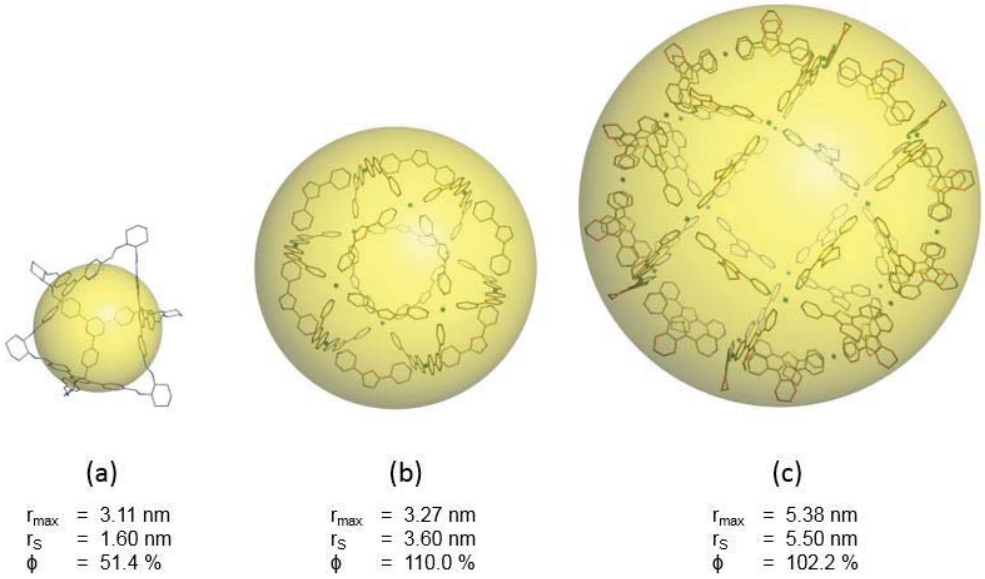


Figure 2.10. Molecular models of (a) "Cage 1" (ref. 37), (b) "$M_{12}L_{24}$" (ref. 38) and (c) "$M_{24}L_{48}$" (ref. 39). Yellow spheres represent the solvodynamic radius, r_s, that was determined by PFG-NMR, using Equation 3. The maximum radius, r_{max}, of cages in was calculated from crystal structures by shrinking a sphere around each cage until it came into contact with an atom. It is fully defined in Figure 2.18. The size of r_s, relative to r_{max}, is calculated as r_s/r_{max} and shown as ϕ.

2.2 Results and Discussion

PFG-NMR is often used to confirm the structures of covalent cages, but this analysis is not always presented alongside the conditions in which samples were measured (see ref. 18). Thus, in beginning to probe how results are affected by molecular shape, it was important to determine whether experimental parameters, such as viscosity, have a significant impact on D_{obs} for cages of different sizes.

2.2.1 Synthesis of covalent cages for study

A library of known cages was synthesised, which included cages of several topologies consistently found within the literature (Figure 2.11). Many additional cages were supplied by collaborating scientists at the University of Liverpool, and details of their contributions can be found in Table 2.1.

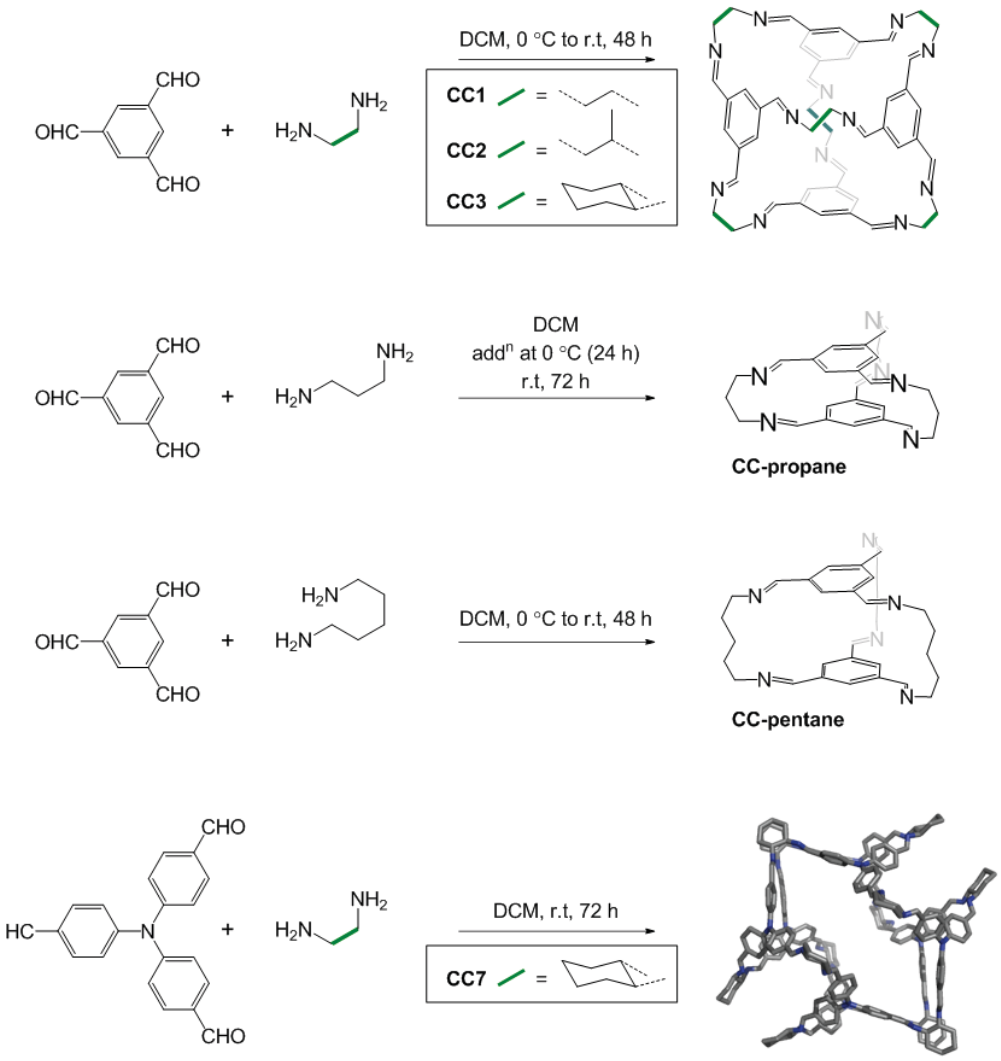


Figure 2.11. Synthetic routes to a library of cages. For simplicity, CC7 has been shown as a capped sticks representation of its single crystal X-ray structure.

Further synthetic steps were taken in the case of **CC1**, **CC3** and **CC7** to alter the functionality and shape of cages in solution.

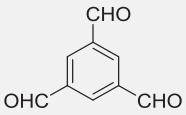
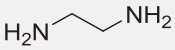
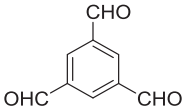
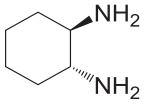
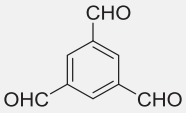
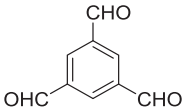
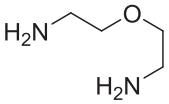
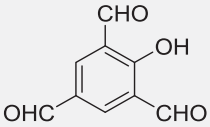
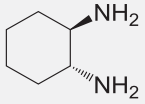
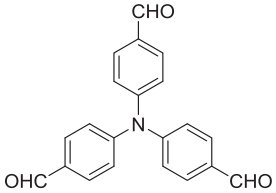
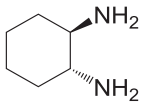
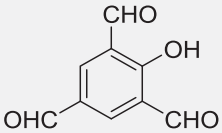
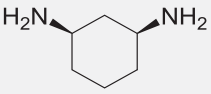
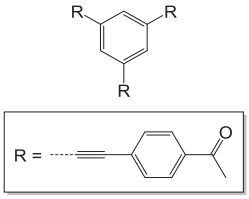
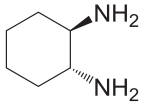
Cage	Aldehyde	Amine	Topology	ref
(R)CC1[a]			4+6	40,41
(R/FT)CC3[a]			4+6	42
CC-propane			2+3	43
Ether cage[b]			2+3	(b)
4OH-CC3[c]			4+6	(c)
(R)CC7[a]			8+12	44
Ico-cage[c]			8+12	(c)
Astetrapod[d]			4+12	(d)

Table 2.1. Cages synthesised for analysis by ^{1}H PFG NMR. (a) In addition to the organic cage, synthetic steps were taken to alter cage structure: reduced cages, denoted by the prefix "R" were also synthesised (See Figure 2.12a). "FT" denotes formaldehyde-tied, as has been reported previously (Figure 2.12b)[42] (b) Cage was synthesised by Heather Armstrong (c) Cage was synthesised by Baiyang Teng (c) Cage was synthesised by Xiaofeng Wu. Topology refers to the ratio of amine to aldehyde molecules in the covalent cage.

CC1 and **CC3** were reduced in the presence of an excess of 24 equivalents of $NaBH_4$ per cage to afford cages connected by amine moieties in place of imines (Figure 2.12a).[41] **CC7** was reduced under the same conditions using 48 equivalents of $NaBH_4$. These cages were expected to be more flexible due to the lack of C=N double bonds, which restrict bond rotation and rigidify the cage.[45] Reduced cages have been reported[42,46] with the prefix "R" and this convention is repeated here. **RCC3** was further functionalised by dissolving it in a solution of paraformaldehyde in methanol and heating until reaction was complete. This afforded a "tied" cage, in which the diamines at each corner of the cage were connected by methylene carbons (Figure 2.12b).[42]

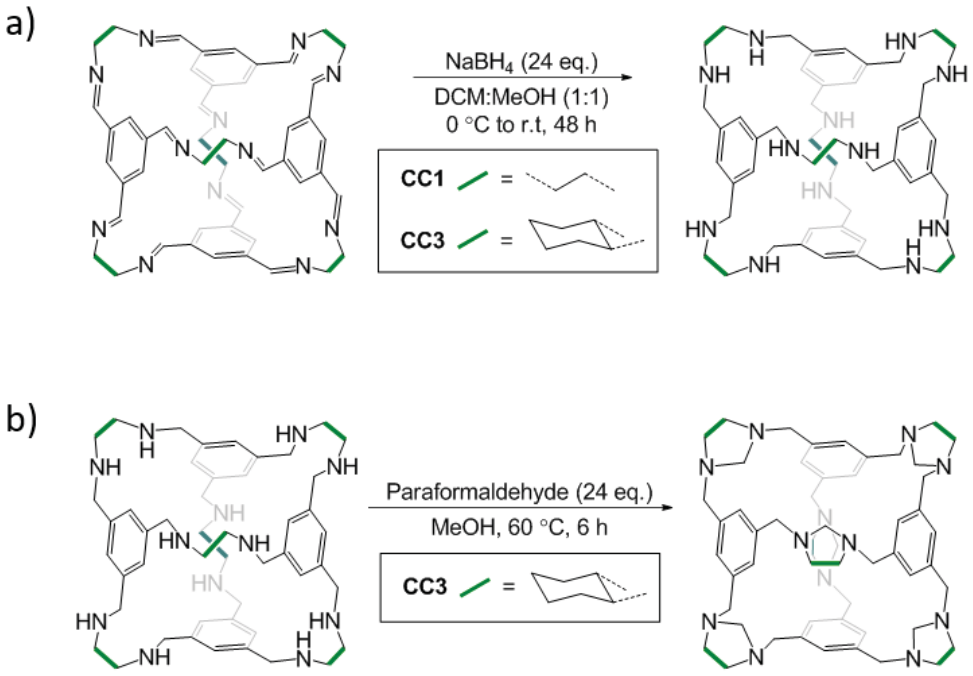


Figure 2.12. (a) Reduction of covalent cages CC1 and CC3 by $NaBH_4$; (b) Post-synthetic functionalisation of RCC3 by paraformaldehyde.

Viscosity has been identified as an important factor in ensuring the validity and reproducibility of PFG-NMR results (Section 2.1.3). Thus, to ensure results were reproducible across a wide range of cage sizes and topologies, the viscosities of several covalent cage solutions were measured at concentrations ranging from 5 to 20 mg mL^{-1} [1] (Figure 2.13). **CC-pentane** (D_{3h}), **CC1** (T_h), and **CC7** (O_h) provided a representative sample of both cage topology, and molecular size. As the molecular weight of cages under study ranged from 523 to 3752 Da, we hypothesised that concentrations measured by mass would provide a more accurate picture of the quantity of material in solution, and the proportional volume occupied by the cages, ϕ (Equation 5, page 6). To confirm this, viscosity was compared with concentrations measured by both metrics. When concentration was measured by mass, increases in viscosity were independent of cage topology (Figure 2.13a), and increased by ~1 % per milligram of cage in a one millilitre solution. Conversely, when concentration was measured by molarity, changes in viscosity were not consistent (Figure 2.13b).

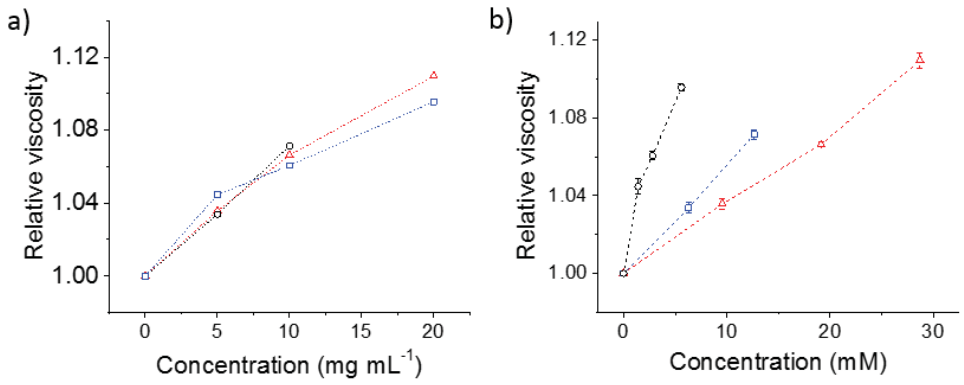


Figure 2.13. A comparison between relative viscosities of cage solutions with concentrations measured by (a) mass and (b) molarity. Cages used: CC-pentane (red triangles), CC1 (blue squares) and CC7 (black circles). Samples were measured in chloroform, and values were normalised by dividing measured viscosities by the native solvent viscosity, 0.542 cP. For clarity, error bars have been omitted in (a), but are identical to those in (b).

To test whether changes in concentrations correlated with D_{obs} in addition to viscosity (Equation 5), ^{1}H PFG-NMR was conducted on the library of cages listed in Table 2.1, at solutions ranging from 1 to 10 mg mL^{-1}. In some cases, concentrations were limited by solubility. Solvodynamic radii were calculated using the Stokes-Einstein equation (Equation 3, radii are presented in Figure 2.14a), and D_0 was calculated as in Figure 2.5. The solvodynamic radius at infinite dilution, $r_{s,0}$, was calculated from D_0 and solvodynamic radii were normalised by dividing by $r_{s,0}$ to provide a measure of change that was independent of size (Figure 2.14b – d).

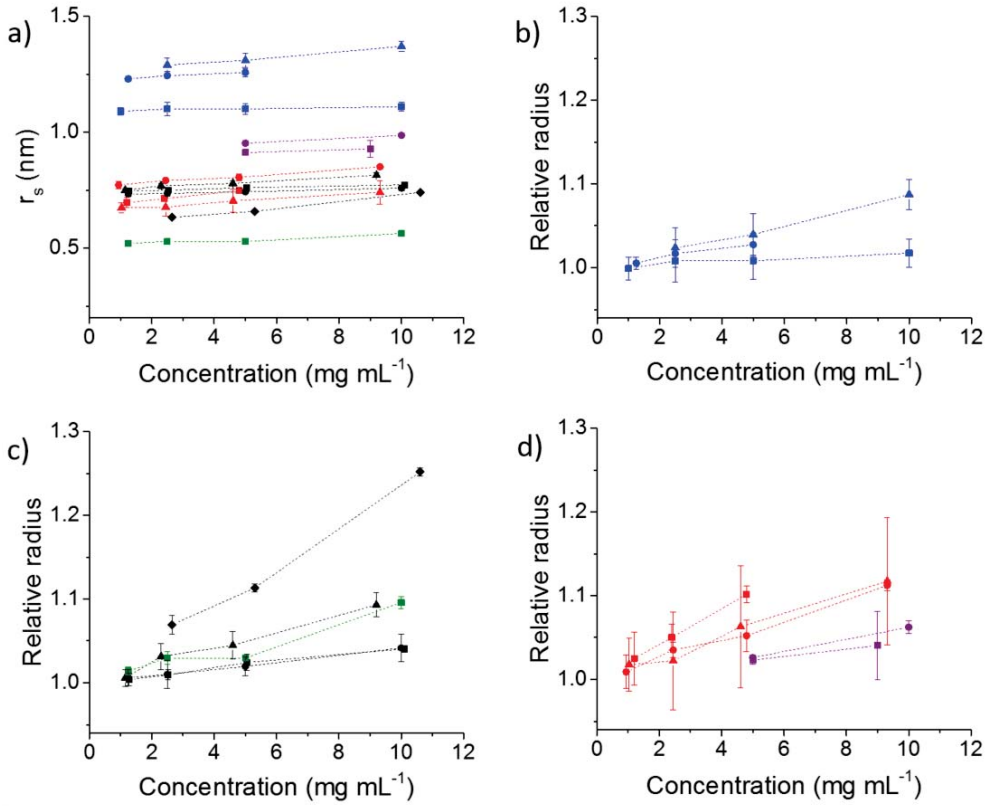


Figure 2.14. (a) Solvodynamic radii of cages at varying concentrations. In some cases, error bars are smaller than the data point size (b–d) Solvodynamic radii of cages normalised by $r_{s,0}$ at varying concentration; Cages are colour-coded by size and type. Organic cages: Astetrapod (blue triangles), CC7 (blue circles), Ico-cage (blue squares), CC1 (black diamonds), CC3 (black triangles), FT-RCC3 (black circles), 4OH-CC3 (black squares), ether cage (green squares). Reduced cages: RCC3 (red) was measured in three solvents: methanol (circles), chloroform (squares) and toluene (triangles). RCC1 (mauve) was measured in two solvents: D_2O (circles), methanol (squares). Where solubility of samples was limited, diffusion coefficients were measured within a more limited concentration range.

In all cases, r_s was found to increase by approximately 1% per mg of cage in a one mL solution (Figure 2.14b–c), in line with viscosity measurements. This indicated that solute-solute interactions had minimal impact on the apparent size of the cage, because changes in D_{obs} were a result of changes in viscosity, and not increasing ϕ. To confirm this, diffusion coefficients of **RCC1** and **RCC3** were measured in a variety of solvents. By varying the polarity of the solvent, the strength of cage-cage interactions could be probed, as the degree of solvation was expected to change.[47,48]

RCC3 appeared to swell between solvents, increasing in solvodynamic radius from 0.67 nm in toluene to 0.77 nm in methanol (Figure 2.14a). 0.77 nm was consistent with the size of its imine parent cage, **CC3**, which had a solvodynamic radius of 0.75 nm. On this basis, we hypothesised that **RCC3** may collapse in less polar solvents. In future work, this could be further investigated by comparing the predicted sizes of cages in different solvents, which has been accomplished with **CC7**,[49] and by varying the polarity of solvents more systematically.

We attributed the slight increase in r_s between **CC3** and **RCC3** in methanol to hydrogen-bonding between the amine nitrogen atoms in reduced cages and solvent molecules in solution (see Section 2.1.1.). This was more apparent with **RCC1**, where r_s increased from 0.63 nm to 0.92 nm between **CC1** and **RCC1**. Data for **CC1** was collected in chloroform, while data for **RCC1** was collected in methanol, and, although we do not anticipate the calculated size of **CC1** to vary between solvents, it would be useful to confirm this experimentally. The effects of solvent are investigated in Section 3.2.1.1., although we do not measure **CC1** in methanol. Furthermore, r_s increased from its value of 0.92 nm in methanol to 0.97 nm when measurements were conducted in water, which is more polar than methanol. This indicated solvation was the most likely cause of the increase in r_s between imine and amine cages.

The rate at which D_{obs} decreased with concentration did not appear to depend on concentration (Figure 2.14d). Significant increases in both the size and solvation of cages, which would be expected to increase *k* (Equation 5) had no statistically significant effect, indicating that contributions to D_{obs} by ϕ were insignificant.

To test the molecular parameter being measured by PFG-NMR in solution, we selected two cages, **CC1** and **CC3**, which are of identical topology, but differ in the degree of steric bulk surrounding the amine nitrogen atoms. As cage synthesis can be

modular, this can be achieved without impacting the basic structure of the cage. This can be seen by highlighting "core" and "external" portions of **CC1** and **CC3** (Figure 2.15).

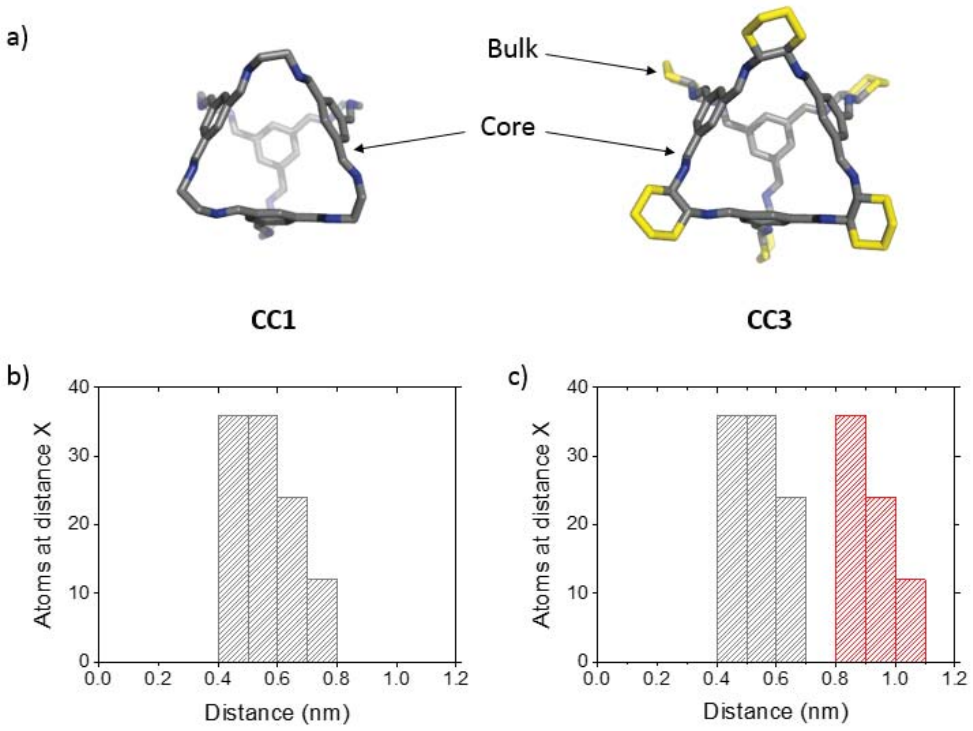


Figure 2.15. (a) Crystal structures of CC1 and CC3, highlighting "core" (grey) and bulky "external" (yellow) portions of the cage. Hydrogen atoms have been omitted for clarity; Histograms showing the number of atoms lying at distances from the cage centre for (b) CC1 and (c) CC3. Grey bars represent atoms within the cage "core". Red bars represent "external" atoms.

Because of the geometry of the cage topology, a significant number of atoms lie within a small radial distribution, which might be considered the "backbone", or core of the cage (Figure 2.15b–c). In the case of **CC1–3**, the structure of the amine is changed to incorporate larger steric functionality, which also impacts solid-state packing.[40] To test whether external bulk effected the observed diffusion coefficient, cages were dissolved in chloroform-*d* at a concentration of 5 mg mL^{-1} and diffusion

coefficients were measured using 16 gradient amplitudes between 1.70 and 32.4 G cm^{-1}. Linear spacing ensured that the Stejskal-Tanner equation could be fitted across the entire gradient range. The diffusion coefficient of **CC1** was found to be 6.46×10^{-10} m^2 s^{-1}, which corresponded to a solvodynamic radius of 0.61 nm. This agreed well with the overall size of the cage molecule as determined by single crystal X-ray crystallography (SCXRD, Figure 2.16a). The solvodynamic radius of **CC3** was calculated to be 1.49 nm, which corresponded with a point mid-way between the core of the covalent cage and the full extent of the cyclohexyl groups, which we refer to as "apexes," as they project into the solvent from the core of the cage (Figure 2.16). This was consistent with the results of Wang et al.,[37] who found the solvodynamic radius of their covalent cage "Cage 1" to be significantly smaller than the full extent of the "apex" groups (see Figure 2.10).

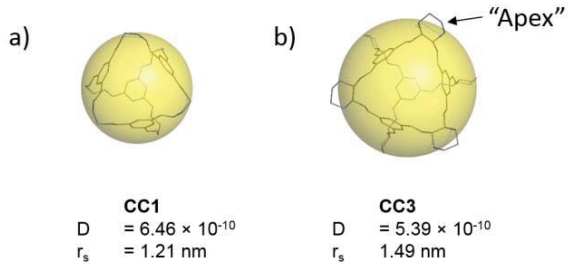


Figure 2.16. Diffusion coefficients and solvodynamic radii of CC1 and CC3. Yellow spheres correspond to the solvodynamic radius as measured by PFG-NMR.

We hypothesised that an increase in the number of external atoms that project into the space surrounding the covalent cage would increase the surface area of the cage that interacts with solvent. This would have the effect of increasing the cross-sectional collision area (Equation 4, page 5), which would increase the measured molecular size. To our knowledge, the extent to which this increase is dependent on the size, shape and separation of "apexes" has not been investigated.

2.2.1.1 Comparing sphericity with PFG-NMR data

Sphericity, Φ_s, is a measure of the degree to which an object's outer surface is smoothly convex. It provides a quantitative framework in which to evaluate surface area increases, which affect collisional cross-sections in solution. Thus, we reasoned that it had the potential to describe the extent to which external functionality should affect PFG-NMR results. It is calculated by comparing the volume (V) of the solid to its surface area (S, Equation 10).

$$\Phi_S = \frac{(\pi)^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{S} \qquad (10)$$

A spherical particle, in which $V = (4\pi r^3)/3$ and $S = 4\pi r^2$, has a sphericity of 1. As materials deviate from this shape, the surface area increases more quickly than the volume, reducing the size of Φ_s (Figure 2.17).

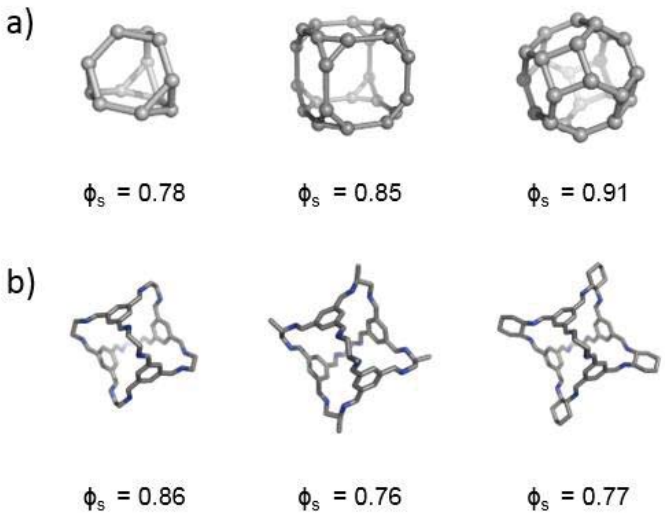


Figure 2.17. Sphericity of (a) common Archimedean solids and (b) covalent cages CC1 (left), CC2 (middle) and CC3 (right). Surface areas and volumes of covalent cages were calculated using ^{3}V Volume Assessor tool,[50] with a probe radius of 5 Å.

It was important to test the degree to which Φ_s could describe the dependence of D_{obs} on structural features that made the molecular surface convex. PFG-NMR data covalent cages in Table 2.1 were compared to their crystal structures, which were downloaded from the crystallographic database service from references listed in Table 2.1. Where solvent was present within the crystal structure, this was removed by hand prior to calculating surface areas and occupied volumes. In the case of the **Ether cage**, **Ico-cage** and **Astetrapod**, where the crystal structure was not known, computational structures were generated by Valentina Santolini at Imperial College London (ICL). Possible conformers were identified using the MacroModel (version 9.9, Schrödinger, LLC, New York, NY, 2011) conformer searching tool, which identified all the low-energy conformers for each structure. All structures within 50 kJ mol^{-1} of the lowest energy structure were retained and further investigated by DFT. If an internal pore was not detected for any of the candidate structures, the process was repeated with an energy acceptance criterion of 500 kJ mol^{-1} in order to see whether any structures with an internal void could be found lying higher in energy. To refine the structures that had been identified using the MacroModel tool, candidate structures were optimised in DFT-D3 calculations performed in CP2K[51] with the PBE functional,[52] TZVP-MOLOPT basis sets,[53] GTH-type pseudopotential,[54] a plane wave grid cutoff of 400 Ry, a cubic box of length 50 Å, and the Grimme-D3 dispersion correction.[55]

We calculate Φ_s for each cage, surface areas and volumes were calculated using ^{3}V Volume Assessor tool.[50] The size of the molecular probe used to generate the sphericity was chosen to reflect the radius of chloroform, in which all diffusion data had been recorded. We estimated the radius of chloroform by averaging its three C-Cl bond lengths (1.75 Å) and its C-H bond length (1.09 Å), and adding the weighted average VdW radius. We calculated the average radius of chloroform to be 3.3 Å, and

selected a probe radius of 5 Å to reflect the repulsive forces felt by molecules at short distances.[56] As the size and topology of each cage was different, it was not possible to compare diffusion coefficients to the sphericity using the crude strategy that had been deployed with **CC1** and **CC3** (see Figure 2.16). Thus, to provide a basis for comparison with PFG-NMR data, two molecular dimensions were extracted from each SCXRD or computational structure (Figure 2.18). The maximum extent of the cage, r_{max}, was calculated by shrinking a spherical mesh, centred about the cage's centre of mass, around the cage until it came into contact with an atom. The weighted average radius, r_{avg}, was calculated using Equation 11:

$$r_{avg} = \sum_{i=1}^{n} V_i . r_i \qquad (11)$$

where V_i is the volume of atom i, r_i is the distance between atom i and the cage's centre of mass, and n is the total number of atoms in the cage.

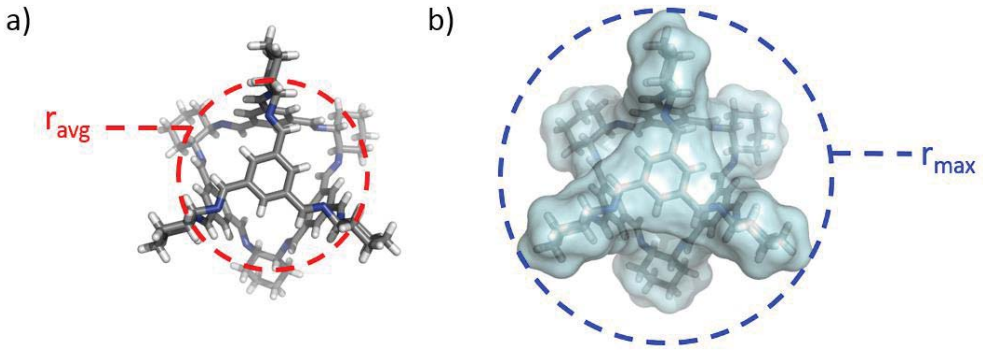


Figure 2.18. (a) The structure of CC3, shown as a capped sticks representation, with r_{avg} shown as a red dotted circle; (b) The structure of CC3 shown as a capped sticks representation, overlaid with the Van der Waals surface of the covalent cage (transparent blue). r_{max} is shown as a blue dotted line.

For consistency, PFG-NMR spectroscopic data was acquired at concentrations of 5 mg mL^{-1}. This data was obtained during the initial comprehensive viscosity testing and is summarised in Figure 2.14a. Initial analysis of this data revealed that molecular

size was the most important factor in determining topology within this data set. Within the series of cages studied, cages with a [3+2] topology were found to have a solvodynamic radius, r_s < 0.6 nm, whilst tetrahedral [4+6] cages were found in the region 0.6 nm < r_s < 0.85 nm. Cages with [8+12] topology had larger solvodynamic radii, above 1.1 nm. Though this is also dependant on the size of the starting materials, this provided the basis for implementing PFG-NMR as a high-throughput screening tool, which is discussed in Section 2.2.2.

Measured solvodynamic radii were compared to the extracted molecular dimensions r_{max} and r_{avg}, generating a qualitative comparison between experimental PFG-NMR spectroscopic data, and molecular size. From this analysis, we observed that more spherical cages such as **CC1** generated a measured solvodynamic radius that lay close to r_{avg}. However, large 'stellated' cages produced solvodynamic radii closer to r_{max} (Figure 2.19). In the most extreme case, **Astetrapod** generated a solvodynamic radius of 1.31 nm, which was equidistant between r_{avg} and r_{max}.

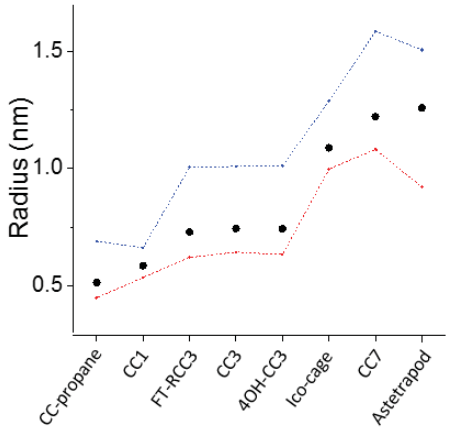


Figure 2.19. A comparison between r_s (black circles) and molecular dimensions calculated from X-Ray crystal structures r_{max} (blue) and r_{avg} (red). Dotted lines are intended to guide the eye.

To determine how r_s varied between r_{avg} and r_{max}, we defined the position-parameter ρ_r, which expressed the NMR-determined radius as a fraction of the distance between the extracted molecular parameters (Equation 12).

$$\rho_r = \frac{r_{NMR} - r_{avg}}{r_{max} - r_{avg}} \quad (12)$$

To test whether this quantification could be used to explain the disparity between observed solvodynamic radii and calculated molecular sizes, the diffusion coefficients and molecular structures of macromolecular cages in Table 2.1 were used to calculate ρ_r. These were compared with sphericities calculated using the using ^{3}V Volume Assessor tool.[50] As expected, an inverse correlation was demonstrated between ρ_r and Φ_s (Figure 2.20), which confirmed our initial hypothesis that more incurvate cages would generate a disproportionally small diffusion coefficient, and so appear larger by PFG-NMR. Larger cages, with [8+12] topologies such as **CC7** and **Ico-cage**, were observed to have higher ρ_r values, which corresponded to the increased degree of incurvature in the surface. The value of ρ_r was observed to vary from 0.23 to 0.31 between **CC1** and **Ico-cage**.

One anomaly was noted in this series, as **Astetrapod** was found to have a significantly higher ρ_r value of 0.58. **Astetrapod** adopts a [4+4] topology (Figure 2.20b), similar to that of **CC11** and **CC12** published by Briggs et al. previously.[57] This topology is referred to as "tetrapodal", and is similar to that of a caltrop (Figure 2.20b, inset), with four apexes protruding from the centre of the molecule. The distance between the ends of these apexes is 1.6 nm (Figure 2.20b), with each apex extending 1.2 nm outwards from the cage "core". We reasoned that, as these apexes are also wedge-shaped, and expand as they protrude into space, the effect may be multiplied by increasing the density of atoms at distances far from the centre of the cage. **CC7**, which also adopts a stellated shape (see Figure 2.11), consists of apexes that narrow

as they project into space, and have an apex-apex distance of 4.6 Å. Thus, we reasoned that the solvent probe would be unable to penetrate the incurvature of **CC7** as deeply as it could in **Astetrapod**, because the size of the incurvature in **CC7** was smaller, being comparable with the size of the solvent. We hypothesised that, by altering the size of the solvent molecule, it may be possible to probe the degree of incurvature present in the molecular surface of the covalent cage.

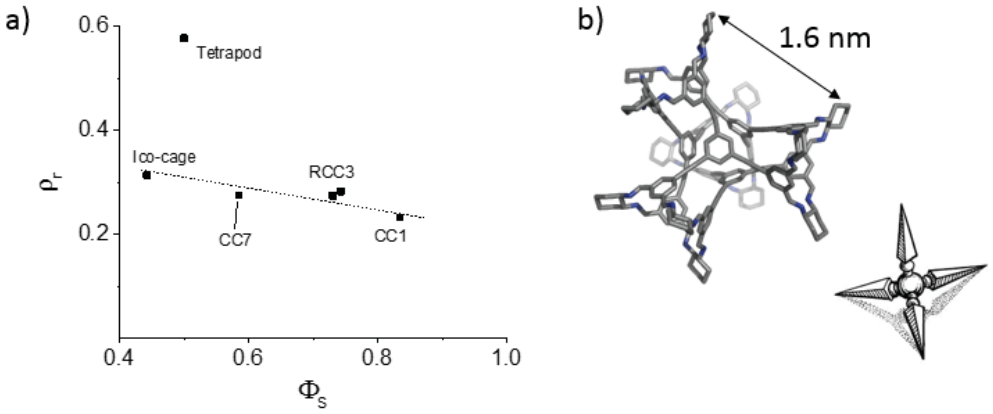


Figure 2.20. (a) The correlation between sphericity (Φ_s) and ρ_r. (b) The computationally-predicted structure of Astetrapod. The distance between cyclohexane groups on adjacent "apexes" is highlighted. Hydrogen atoms have been omitted for clarity. The image of a caltrop (inset) has been included as a comparison of shape. Image taken from the public domain.

2.2.1.2 Using solvent probes to examine molecular shape

As solvents differ in size, they also differ in their ability to access the external space around molecules with which they interact. This is affected not only by solvent size, but also by polarity, and solvent-solute interactions such as H-bonding. In cases where the external shape of the molecule is highly incurvate, a small molecule may be able to pass more closely to the molecule without interacting with it. Empirically, this would result in the molecule *appearing* smaller from the point of view of the probe

molecule. The Connolly surface area,[58] which represents the parts of the molecule that can interact with solvent, can be used to visualise this (Figure 2.21)

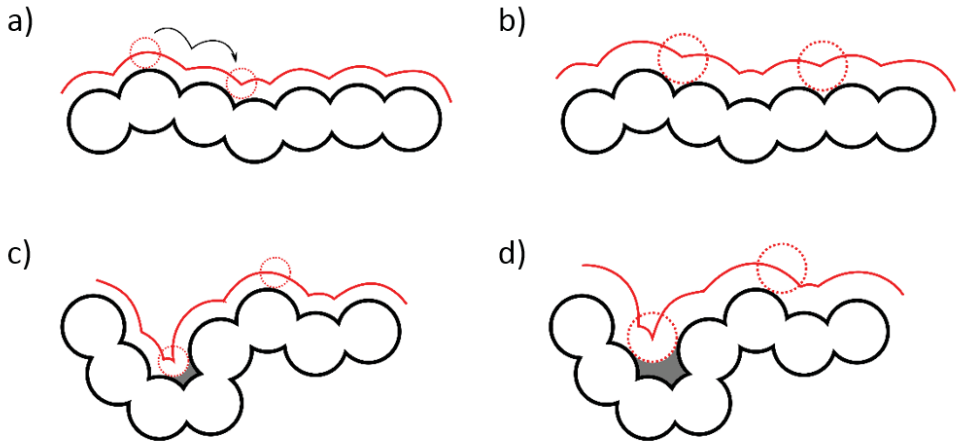


Figure 2.21. A schematic representation of Connolly surface areas for two types of solids; (a) and (b) depict a molecule with minor incurvature; (c) And (d) depict a molecule with increased incurvature. Dotted circles represent the size of the solvent molecule. The black arrow in (a) is intended to represent rolling the solvent across the molecular surface, which is the process by which surface areas are calculated.

We hypothesised that, in principle, altering the molecular radius of the solvent would change the average collision radius (σ_{avg}, Section 2.2.1) in cases where cages were not convex. We predicted that altering the size of the probe solvent with prismatic cages, where incurvature is limited (Figure 2.22a), would have a small effect on σ_{avg}. However, this effect was expected to increase as the geometry of the covalent cage changed to tetrahedral (**CC3**, Figure 2.22b) and cubic (**CC7**, Figure 2.22c).

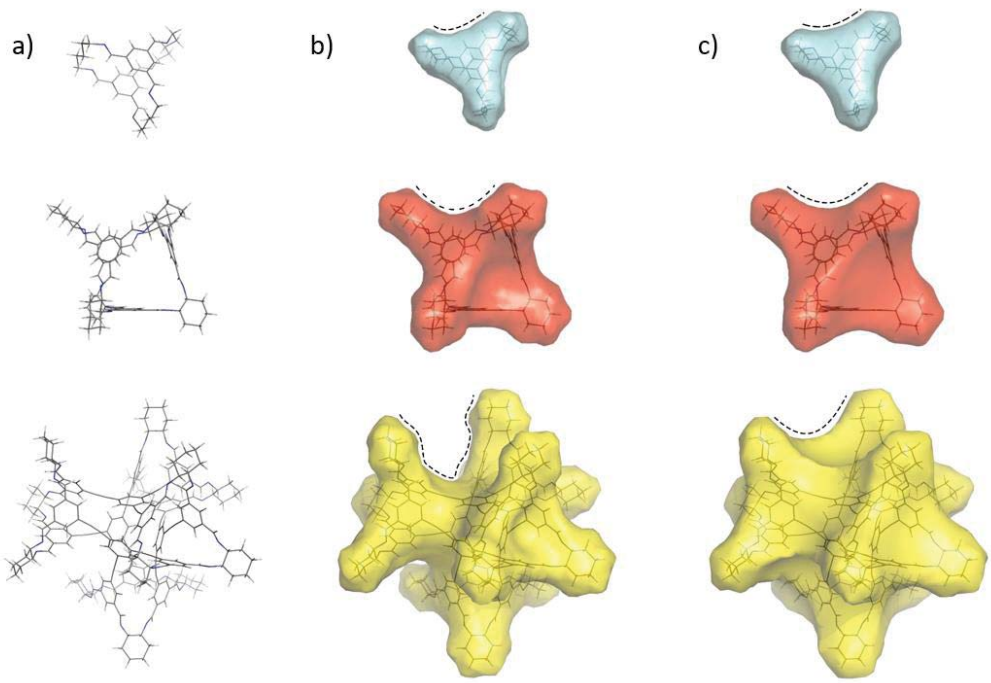


Figure 2.22. (a) X-ray crystal structures of CC-pentane (top), CC3 (middle) and computationally-generated structure of Astetrapod (bottom); Connolly surface areas of cages in (a) with probe radii of (b) 2.4 Å and (c) 6.4 Å. Dotted lines are intended to highlight concavity in the external surface. Probe radii were selected to be larger than window diameter.

To test this hypothesis, solutions of cages **CC-pentane**, **CC3**, and **CC7** were prepared in two solvents of comparable polarity and viscosity, but with a systematic increase in solvent radius: dichloromethane-d_2, with an estimated radius of 2.9 Å, and chloroform-*d*. Cages were dissolved at a concentration of 5 mg mL^{-1}, to remove any concentration effects from the results obtained, and diffusion NMR measurements were obtained using a double-stimulated emission pulse sequence (dstegp3s)[59,60] to remove the effects of convection. Diffusion coefficients were converted to solvodynamic radii using the Stokes-Einstein equation (Equation 3).

Solvodynamic radii were observed to increase on changing the solvent from dichloromethane-d_2 to chloroform-*d*. However, the proportional increase appeared to dramatically increase from **CC-pentane** and **CC3** to **CC7** (Figure 2.23). **CC7** was

found to increase in apparent size by 17 % between dichloromethane-d_2 and chloroform-*d*, while **CC-pentane** and **CC3** were found to increase by 7.6 and 7.7 % respectively.

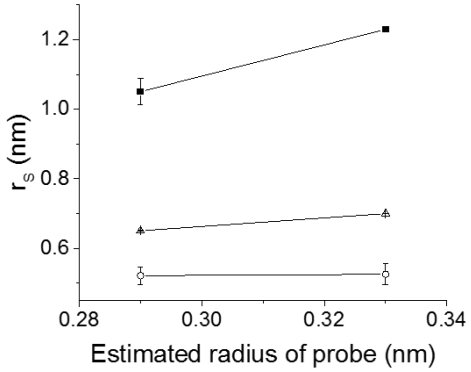


Figure 2.23. Solvodynamic radii of CC-pentane (empty circles), CC3 (empty triangles) and CC7 (filled squares) in chloroform-*d* and dichloromethane-d_2.

We tentatively attributed this difference to the increase in incurvature between smaller cages and **CC7**. However, we were unable to corroborate our findings by analysing a larger range of cages. Given constraints on the quantity of material of **Astetrapod** and **Ico-cage**, it was not possible to obtain data in dichloromethane-d_2, which would have helped provide a more detailed picture of the effects of apex shape and distribution. Given adequate time to expand these results, it would be useful to conduct testing on a wider range of cage topologies, which would allow a comprehensive analysis of covalent cage shape to be undertaken.

2.2.1.3 Interpreting PFG-NMR results on the basis of analogous cage molecules

Changes in cage structure have often been found to arise from small structural changes in the starting materials used.[38,39] One example of this is to increase the length

of aliphatic terminal diamines in a reaction with benzene-1,3,5-tricarboxaldehyde (TFB). The first diamine in this series, ethylenediamine, reacts with TFB to form **CC1**. Subsequently, increasing the number of methylene carbons between the nitrogen atoms of the diamino starting material generates two distinct series of cages.[43] Diamines with an even number of methylene carbons between the nitrogen atoms can adopt a gauche conformation, which favours the formation of tetrahedral cages (Figure 2.24a). Conversely, where the number of methylene carbons is odd, a gauche configuration is no longer possible and diamines form prismatic cages (Figure 2.24b).

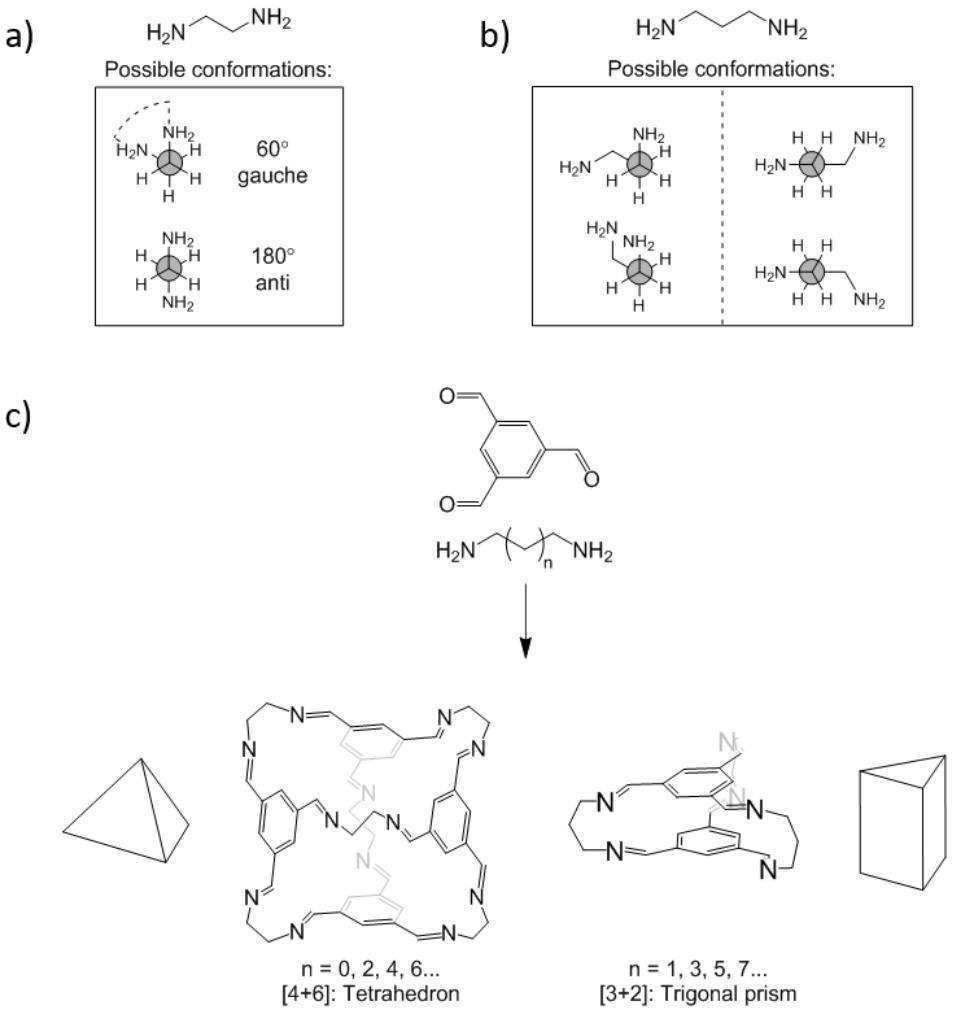


Figure 2.24. Geometry of (a) even- and (b) odd-length diamines; (c) the reaction between diamines and TFB affords cages of different topologies.

In combining these diamines with TFB, we expected two sets of materials to be formed (Figure 2.24c). As with **CC1**, "even" diamines, which are stable in gauche conformations, were hypothesised to form tetrahedral cages and odd linear diamines were predicted for form cylindrical cages. Calculations performed by Kim Jelfs confirmed this, indicating that a sizeable energy gap existed between cage topologies in each case.[43]

Syntheses were optimised in order to both maximise yield and afford material for characterisation. In the case of 1,3-propanediamine, 1,5-pentanediamine and 1,7-heptanediamine, slow addition of the diamine over 8 to 24 hours was sufficient to inhibit precipitation of polymeric material. However, with 1,4-butanediamine and 1,6-hexanediamine, a slower rate of addition was required. Thus, prior to determining the topological form of each product, noticeable synthetic differences between "odd" and "even" products (where odd and even refer to the number of methylene carbons between vicinal diamine nitrogen atoms) were noted. We reasoned that this may indicate that different reaction products were being formed. Furthermore, we hypothesised that this may have resulted from the increase in complexity between prismatic to tetrahedral cages, although this is impossible to determine by one-dimensional ^{1}H NMR spectroscopic analysis. Optimised conditions are listed in Table 2.2. Cages were named according to the length of the methylene carbon chain between nitrogen atoms in the aliphatic starting diamine. Thus, 1,3-diaminopropane reacted with TFB to form **CC-propane**.

Cage	Addition temperature (°C)	Addition time (h)	Mass (m/z)	Topology[a]
CC-propane	0 °C to r.t.	24 h	439.3	[3+2]
CC-butane	0 °C	24 h	961.6	[6+4]
CC-pentane	r.t.	6 h	524.5	[3+2]
CC-hexane	0 °C to r.t.	24 h	1129.8	[6+4]
CC-heptane	0 °C to r.t.	6 h	607.4	[3+2]
CC-octane	0 °C[a]	48 h	649.5	[3+2]
CC-nonane	0 °C to r.t.	12 h	691.5	[3+2]

Table 2.2. Optimised conditions for the synthesis of organic covalent cages. (a) Topology of cages was determined by PFG-NMR and mass spectroscopic analysis and is discussed later in this Section.

The stability of "odd" products was noted to increase from **CC-propane** to **CC-pentane**, with **CC-propane** being unstable to isolation. We attempted to purify the material by solvent evaporation and by using a solvent-swap to methanol, although both processes afforded an insoluble white precipitate. Deliberate precipitation into hexane was found to afford insoluble material, which we attributed to polymerisation. Finally, attempts were made to crystallise the material slowly. In this case, crude reaction mixtures were diluted into dichloromethane to a concentration of 1 mg mL^{-1}, and placed in 4 mL vials. These were placed into 12 mL vials to which was added a range of organic solvents: MeOH, EtOAc, $CHCl_3$, IPA, Et_2O, *i*Pr_2O, EtOH, acetone, THF, hexane, MeCN and Toluene. Where solids were formed, these were found to be insoluble in common organic solvents. Reactions between TFB and 1,3-

diaminopropane produced an off-white precipitate, which we believed to be polymeric material on the basis of its insolubility, and generated a small quantity of solution-stable covalent cage. To determine the conversion of starting materials to **CC-propane**, ~ 5 μL of dimethylformamide (DMF) was added to the initial solution of 1,3-diaminopropane prior to reaction, and the concentration of DMF was calculated by ^{1}H NMR spectroscopic analysis of the starting solution. The quantity of cage in solution was calculated by comparing the integral area of chemical shifts assigned to **CC-propane** with that of DMF at 2.71 ppm, and was found to be 24 %.

CC-pentane was found to be partially stable and, upon isolation, a diminishing quantity of **CC-pentane** could be recovered by dissolution. This indicated that **CC-pentane** is stable in solution, but begins to degrade on isolation. Attempts to characterise this material by elemental analysis were found to be impossible. CHN values appeared consistent with material that had been partially hydrolysed, with hydrogen content higher than expected, and carbon and nitrogen content lower (Table 2.3). **CC-heptane** and **CC-nonane** were also found to be stable in solution, but degraded on isolation as a solid compound.

Element	**Experimental**	**Theory**	
		CC-Propane	CC-propane·(0.7)H_2O
C	74.07	75.82	74.04
H	8.08	8.10	8.17
N	15.66	16.08	15.7

Table 2.3. Elemental analysis of CC-propane samples after isolation

Conversely, in the case of even-numbered diamines, stability appeared to drop with increasing chain length. Stability of "even" products was found to decrease from **CC1** to **CC-butane**. Reactions between TFB and 1,4-diaminobutane produced an off-white precipitate, which we believed to be polymeric material on the basis of its insolubility, and conversion to **CC-butane** was calculated using a method identical to that used with **CC-propane**. Conversion was found to be 27 %. ^{1}H NMR spectroscopic and mass spectroscopic analysis of reactions between TFB and 1,6-diaminohexane demonstrated a large number of chemical shifts, which we initially attributed to a mixture of intermediates and topologies being formed (Figure 2.25). The presence of downfield chemical shifts at ~10 ppm indicated the presence of unreacted aldehyde moieties, which was indicative of partially-reacted starting material, and oligomeric intermediates. Furthermore, both [3+2] and [6+4] topological cages were identified in the mass spectra of crude reaction mixtures (Figure 2.25b and c).

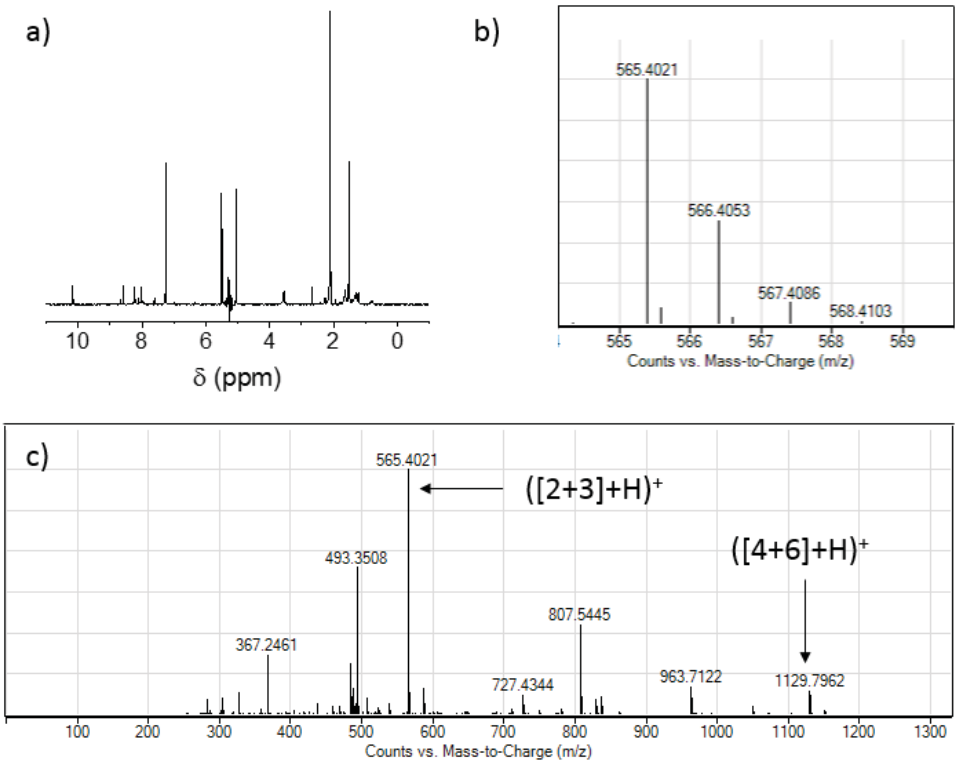


Figure 2.25. (a) ^{1}H NMR spectrum of a crude reaction between TFB and 1,6-diaminohexane, with a soft presaturation pulse centred around 5.29 ppm; (b) high resolution mass spectrum of the same crude reaction, focussed around 567 Da; (c) Full mass spectrum of the same crude mixture.

The quantity of each intermediate and topology was difficult to measure by ^{1}H NMR spectroscopic analysis, as [2+3] and [4+6] cages do not have indicative chemical shifts that allow them to be differentiated. However, we hypothesised that diffusion NMR could be used to analyse the size of molecules in solution and identify which chemical shifts corresponded to which particle. Attempts to isolate **CC-hexane** in either topological form, such as by precipitating material or removing solvent, produced an intractable off-white solid, which we attributed to the formation of polymeric material. Thus, in order to facilitate *in situ* analysis, it was important to develop a solution that would allow ^{1}H PFG-NMR of reaction mixtures to be carried out without isolation. Furthermore, because ρ_r was known to vary with cage topology,

PFG-NMR analysis was conducted on the series from two to seven methylene carbons, which would allow diffusion coefficients to be compared between cages in the series.

Crude reaction samples in dichloromethane were diluted into $CDCl_3$ to facilitate deuterium lock, and a soft pre-saturation pulse was applied prior to both one-dimensional ^1H NMR pulse sequences, and ^1H PFG-NMR. This was focussed around the centre of the DCM peak at ~5.30 ppm and acted to suppress signals around solvent chemical shift. As viscosity is an important parameter in PFG-NMR evaluation (see Section 2.1.3), it was important to evaluate the effect of diluting DCM reaction mixtures into $CDCl_3$, as solvent mixtures can display non-uniform viscometry characteristics upon mixing. To investigate the effect of dilution, **CC1** was dissolved in solvent mixtures from 100:0 to 0:100 $CDCl_3$:DCM in 20 % increments, and diffusion NMR performed on each mixture, with an additional mixture where solvent quantities were equal. A linear relationship was observed between the volume fraction of $CDCl_3$, χ_{CDCl_3}, and both the diffusion coefficient (Figure 2.26a) and viscosity of dilute solutions (Figure 2.26b). This indicated that solvents behaved ideally on mixing. Furthermore, this allows η to be calculated from χ_{CDCl3} at the time of analysis, removing the need to measure viscosities separately.

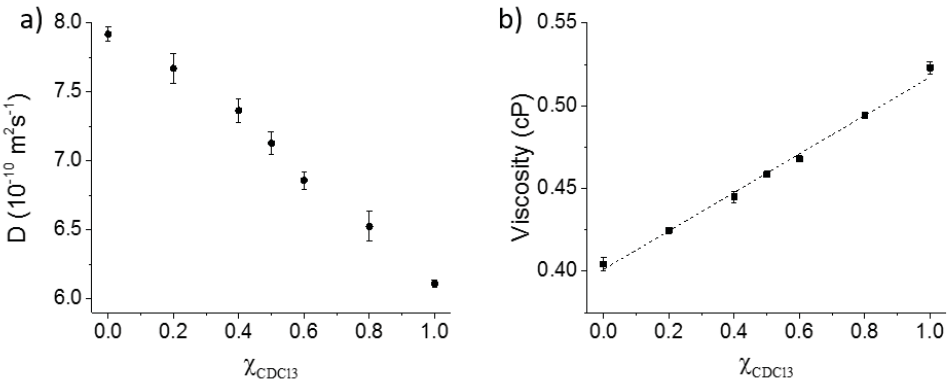


Figure 2.26. (a) Diffusion coefficients obtained by ^{1}H PFG-NMR spectrometry and (b) viscosities of DCM:$CDCl_3$ mixtures containing CC1.

Using optimised conditions, each reaction was sampled upon completion and crude solutions were diluted to form 1:1 mixtures with $CDCl_3$. It was assumed that the viscosities were unchanged by the presence of small quantities of other material, which was supported by previous testing (see Section 2.2.1, Figure 2.14).

As described above, PFG-NMR was performed using a modified pulse sequence, which included a pre-saturation pulse to suppress solvent signals at 5.5 ppm. A comparison of amine chain length with diffusion coefficients revealed two starkly different sets of materials (Figure 2.27). The observed diffusion coefficient of **CC-hexane**, 5.18×10^{-10} m^2 s^{-2}, which was consistent with the size of tetrahedral cages with [4+6] topology. Furthermore, the diffusion coefficient of **CC-hexane** was smaller than the diffusion coefficient of **CC-heptane**, which was found to be 5.25×10^{-10} m^2 s^{-2}. This indicated **CC-hexane** existed as a larger species than **CC-heptane**, which we reasoned was unlikely if the cages shared a common topology. However, one explanation for the difference in size could be structural collapse of the covalent cages, in which cages collapse in on themselves, losing their structural rigidity. Thus, if **CC-heptane** adopted a prismatic geometry, but collapsed, it might be expected to have a

smaller solvodynamic radius than a prismatic **CC-hexane**. To further elucidate the structural features of these cages, each series was investigated using ρ_r as a benchmark by which to compare the results obtained. This was combined with molecular modelling, completed by Kim Jelfs at ICL which looked at the structural persistence of cages in solution and attempted to estimate whether cages were likely to collapse in on themselves.

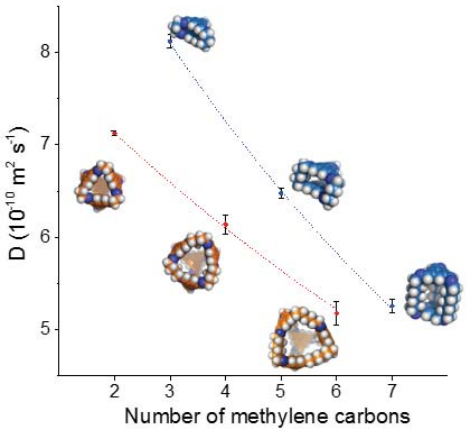


Figure 2.27. ^{1}H PFG-NMR spectroscopy reveals two sets of materials: tetrahedral cages (red) and prismatic cages (blue) were apparent from the different rates at which sizes were observed to increase. Dotted lines are intended to guide the eye.

2.2.1.4 "Odd cages"

In attempting to produce ρ_r values for prismatic cages in the "odd" series, it was important to identify the most probable shape of each covalent cage in solution prior to comparison. Molecular models of each covalent cage in the series were produced by Kim Jelfs (ICL), and the structural persistence of each cage was evaluated to determine whether it was possible for the cage to collapse into itself, and whether this represented an energetic advantage. Molecular modelling suggested that structural collapse was possible at higher chain lengths. Energy-minimised structures indicated that **CC-pentane** would be structurally persistent, but that **CC-heptane** and **CC-**

nonane would collapse. This was attributed to the increased flexibility and conformational freedom, and this indicated that it may be possible for the PFG-NMR data to be consistent with a [3+2] topology for **CC-hexane**.

In the case of **CC-heptane** and **CC-nonane**, two molecular models were produced. An energy-minimised structure, which was likely to represent the average molecular structure in solution (**CC-heptane** is shown in Figure 2.28a), is referred to as the 'collapsed' structure, as the prismatic geometry has collapsed in on itself to produce the equilibrium conformer. A second, structure was generated, which we refer to as 'inflated', which corresponds to the lowest energy conformer in which a permanent, structurally-persistent cavity could be found (Figure 2.28b).

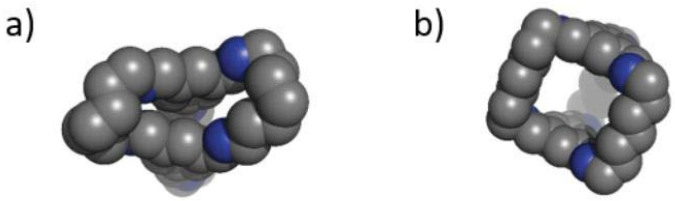


Figure 2.28. Computationally-generated (a) inflated and (b) energy-minimised structures of CC-heptane.

Prior to producing values of ρ_r, these computational structures were used to generate diffusion coefficients from Equations 7–9. For prismatic cages, two parameters were measured in addition to r_{max} and r_{avg}. The axial distance (*l*) was calculated by measuring the distance between each phenyl ring, and the diameter, *d*, was measured by viewing the cage along the arene-arene axis and shrinking a circle around the cage until it came into contact with an atom (Figure 2.29b).

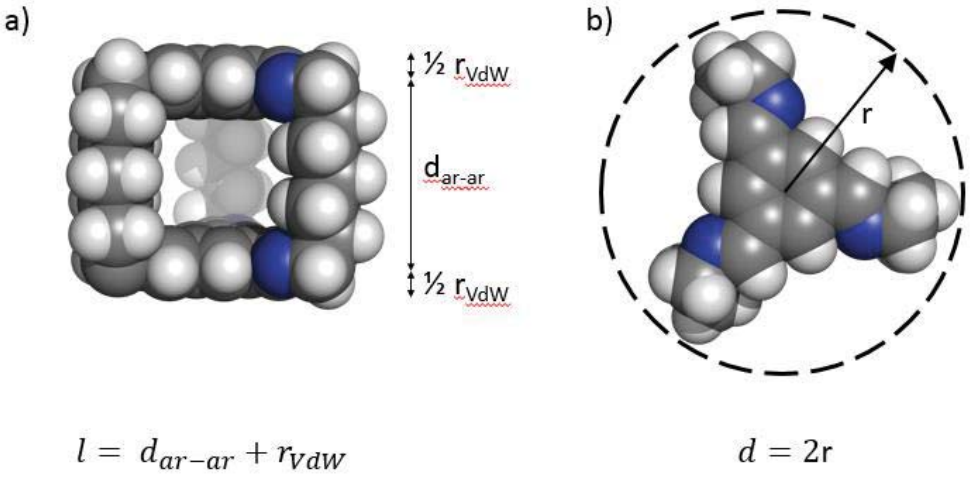


Figure 2.29. (a) The axial dimension of prismatic cages was calculated by measuring the axial-axial distance; (b) The diameter was calculated by viewing the cage along the arene-arene axis and shrinking a circle about the cage until it came into contact with an atom.

Equations 3, 8 and 9 were used to generate diffusion coefficients which might be expected for each cage. **CC-propane**, where $l < d$, was modelled as an oblate spheroid (Equation 9). **CC-pentane** ($l \approx d$) was modelled as a sphere (Equation 3). Larger prismatic cages, where $l > d$, were modelled as prolate spheroids (Equation 8), with l and d as the major and minor axes respectively (Figure 2.30).

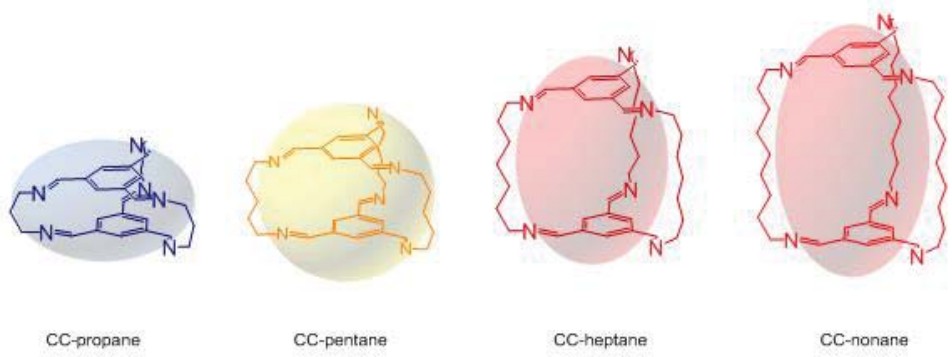


Figure 2.30. Schematic examples of "odd" products. Representations of anisotropic shapes used to model their geometry are overlaid; (a) CC-propane, with an oblate spheroid (blue); (c) CC-pentane with a sphere (yellow); (d) CC-heptane and CC-nonane with prolate spheroids (red).

Cages that had been predicted to remain structurally persistent in solution, **CC-propane** and **CC-pentane**, demonstrated excellent agreement was found between inflated structures and experimental data. **CC-Heptane**, which had been predicted to collapse, also appeared to demonstrate good agreement with the spheroidal model. However, **CC-nonane**, which had been computationally predicted to collapse fully, demonstrated poor agreement between PFG-NMR data and a modelled prismatic structure (Figure 2.31). This suggests that **CC-nonane** adopts a collapsed geometry in solution with no persistent pore.

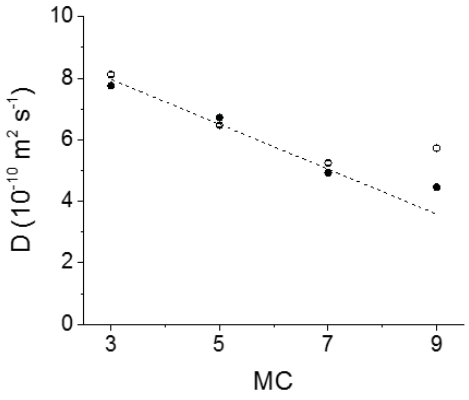


Figure 2.31. Experimental diffusion coefficients (open circles) of prismatic cages in 1:1 mixtures of $CDCl_3$:DCM. Predicted diffusion coefficients (closed circles) were produced from inflated computational structures and provided evidence of collapse where the number of methylene carbons (MC) is greater than five.

Initially, these results highlight the limitations of PFG-NMR in characterising structural details of molecules in solution. In the case of **CC-heptane**, a reasonable agreement was obtained between the prismatic-modelled diffusion coefficient and the experimental one. However, as partial collapse is likely to result in only a small change in D, this was impossible to identify from experimental data. In Section 3.2.1.6, we evaluate these structures using ρ_r.

2.2.1.5 "Even cages"

Within the series of 'even cages', molecular modelling suggested that structural collapse would be energetically favourable in cages larger than **CC-butane**. This suggested that the size of cages in solution may be smaller than expected (Figure 2.32), but also indicated that the geometrical pre-configuration that supports cage formation may deteriorate at longer amine chain lengths.

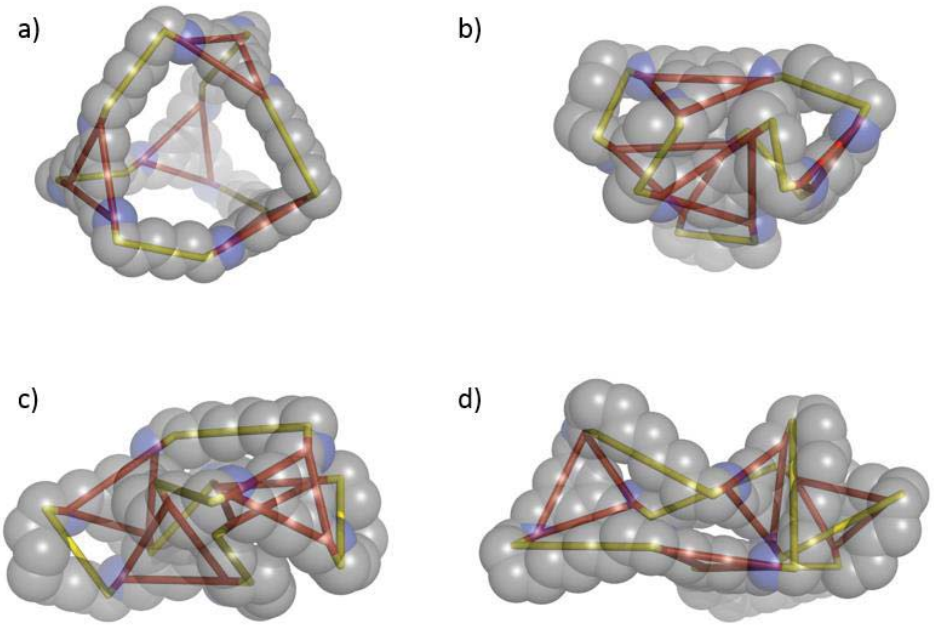


Figure 2.32. Computationally-generated structures of (a) CC-hexane, artificially inflated, (b) CC-hexane, minimised; (c) CC-octane, minimised, and (d) CC-decane, minimised. Hydrogens have been omitted for clarity. Geometrical representation of triamines (red) and diamines (yellow) have been overlaid to demonstrate the loss of symmetry on collapse.

Diffusion coefficients of **CC1**, **CC-butane** and **CC-hexane** were measured *in-situ*, although a large increase in the number of ^{1}H NMR chemical shifts suggested that solution-phase purity prior to isolation began to deteriorate with increasing diamine length (Figure 2.33).

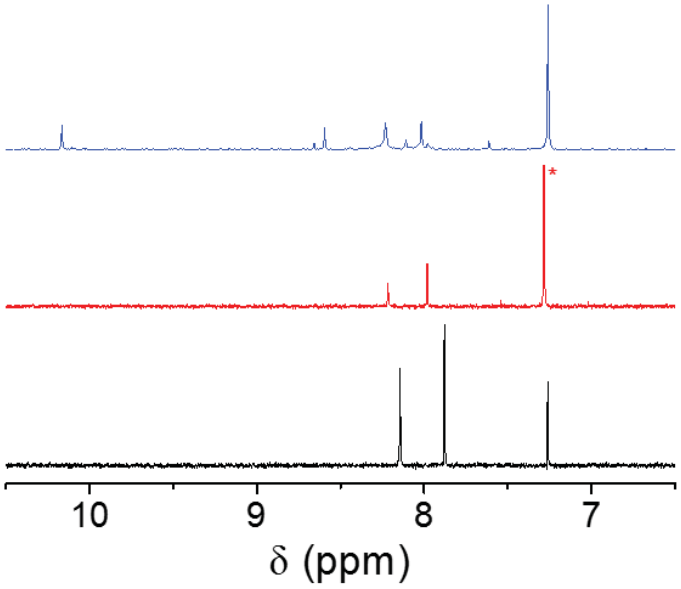


Figure 2.33. ^{1}H NMR spectra of crude reaction mixtures in the synthesis of CC1 (black), CC-butane (red) and CC-hexane. Spectra are focused at the high-field end of the spectrum ($6.5 < \delta < 10.5$) where imine and aromatic protons are present. Chloroform (asterisked) peak in the CC-butane spectrum has been artificially truncated for clarity.

In reactions between benzene-1,3,5-carboxaldehyde and 1,8-daminooctane, slower addition rates generated positive results (Table 2.2). Where the addition of the amine to the aldehyde was conducted in less than 48 hours, PFG-NMR spectroscopic analysis of these crude samples indicated that the size of species in solution ranged from 0.6 to 1.1 nm in diameter, which was not commensurate with macromolecular assembly. Furthermore, a significant quantity of material precipitated over the course of addition. In an example reaction, a solution of benzene-1,3,5-tricarboxaldehyde (65.0 mg, 0.45 mmol, 1.5 eq.) in dichloromethane (50 mL) was added to a cooled solution of 1,8-diaminooctane (49.0 mg, 0.30 mmol, 1 eq.) in dichloromethane (150 mL) dropwise over 24 hours. After addition was complete, a white precipitate was observed to have formed on the glass, which was separated by filtration after 7 days. The resultant white

solid (85 mg) was found to be insoluble in common organic solvents. However, when addition of the amine to the aldehyde was conducted at 0°C over 48 hours, ^{1}H NMR spectroscopic analysis indicated that all aldehyde groups had reacted (Figure 2.34). Furthermore, characteristic[40] chemical shifts at 8 ppm in the ^{1}H NMR spectrum confirmed the presence of imine environments. In spite of this, a significant quantity (75 mg) of precipitated material was recovered when reactions were conducted on the same scale as described above. Attempts to isolate the material by precipitation, solvent-removal and solvent swapping were found to be unsuccessful, and crystallisations produced only insoluble material and intractable films.

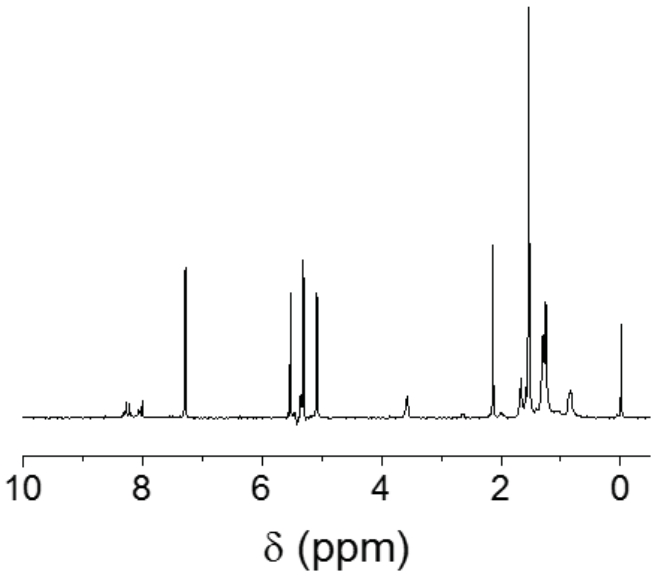


Figure 2.34. ^{1}H NMR spectrum of crude reaction mixture in the reaction of benzene-1,3,5-tricarboxaldehyde with 1,8-dibromooctane.

To identify the topology of **CC-octane**, crude reaction mixtures were analysed by PFG-NMR spectroscopy and mass spectrometry. PFG-NMR spectroscopic analysis

indicated that chemical shifts at 8.19 and 3.58 ppm were associated with a single molecule with a diffusion coefficient of 5.77×10^{-10} m^2 s^{-1}. Because of the limited s/n and peak overlap with the chemical shifts associated with molecules such as water in the sample, it was impossible to confirm this for peaks between 0.5 and 2 ppm. Using Equation 3 (Page 4), the solvodynamic radius was calculated to be 0.83 nm. In comparison to other cages in the series, **CC-octane** appeared most similar in size to collapsed prismatic cages **CC-heptane** and **CC-nonane** (Figure 2.35a).

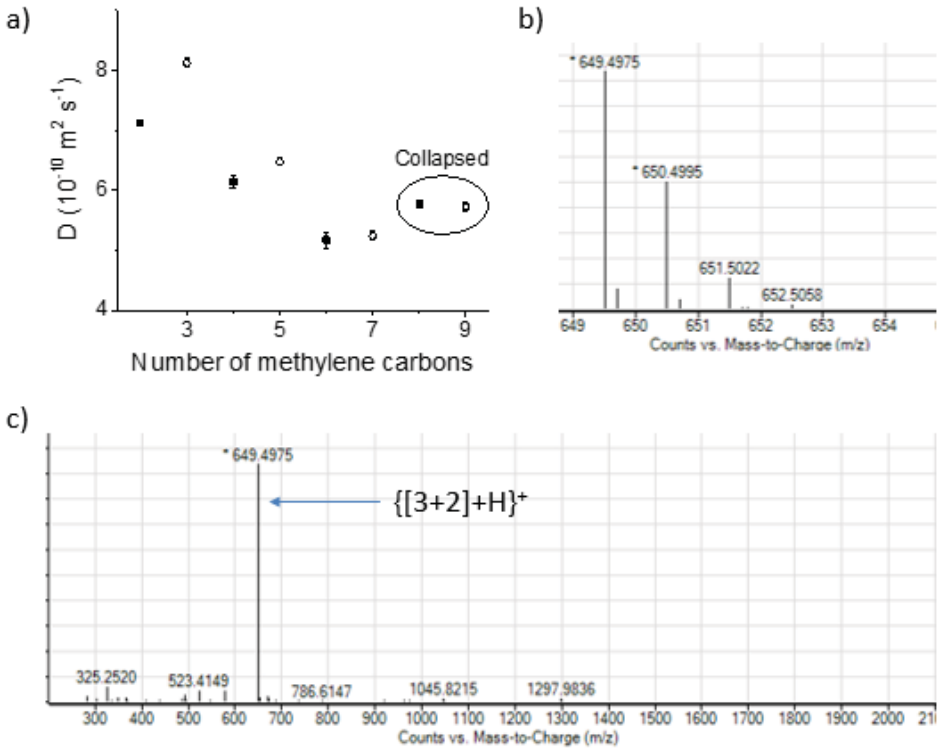


Figure 2.35. (a) Diffusion coefficients of aliphatic cages synthesised by mixing terminal linear diamines and benzene-1,3,5-benzenetricarboxaldehyde. CC-octane and CC-nonane have been circled to highlight similarities in their diffusion coefficients; (b) Mass spectrum of a crude reaction mixture between 1,8-diaminooctane and benzene-1,3,5-tricaroxaldehyde.

Mass spectra of crude reaction mixtures that had been filtered to remove solid material confirmed the presence of **CC-octane** as a cage with [3+2] topology, with

with an experimental mass of 649.4975 *m/z* The expected mass for $C_{42}H_{61}N_6^+$ was 649.4952 *m/z*. Analysis of the isotope pattern in the high resolution mass spectrum (Figure 2.35b) confirmed that this was a singly-charged (z = 1) species. This was the only mass observed in the mass spectrum (Figure 2.35c).

The formation of prismatic geometry in the reaction between 1,8-diaminooctane and benzene-1,3,5-tricarboxaldehyde was attributed to a decrease in the steric and conformational strain found in smaller structures. Templating is crucial in the construction of complex species such as molecular knots and viral capsids.[61,62] In cases like this, intramolecular interactions are often insufficient to favour the formation of discrete molecules in the absence of directional bonding.[63–65] We reasoned that as the diamine chain length increased, the number of conformers available to the amine starting material would increase exponentially: each C-C bond is capable of adopting three conformations, although the energies associated with these are unequal. Thus with chain extension, the number of conformations available increases as 3^x, where *x* is the number of methylene carbons. Furthermore, as the cavity of the product cage becomes larger, the thermodynamic advantage of macromolecular formation is reduced. Instead, the energy gained by VdW interaction of aliphatic chains in the collapsed conformation becomes dominant.

Using PFG-NMR, it is possible to analyse macromolecular cage materials with sufficient accuracy to determine what topology they have adopted. Generally, this is possible without data modelling. However, where the sizes of cages are similar despite topology, a comparison between computational models and experimental NMR data can afford detailed information about cage structure.

Finally, it has been found that details about the incurvature of macromolecular cages can be generated from PFG-NMR spectroscopic analysis, but this requires

several solvents, and experiments. In the following section, PFG-NMR is incorporated into a high-throughput screen, and is used to characterise 27 cage materials *in situ* and without the need for isolation.

2.2.1.6 Values of ρ_r for anisotropic covalent cages

Covalent cages in the homologous series were analysed by comparing their computationally-generated structures with the experimental solvodynamic radii measured by PFG-NMR (Figure 2.36a and b). Values of ρ_r were found to decrease across the even series from two to eight methylene carbons. The value of ρ_r in a structurally-persistent tetrahedral cage, **CC1**, was found to be 0.43, which decreased to 0.23 in **CC-butane**. This increased to 0.25 in **CC-hexane**, although the increase is smaller in size than the error associated with the PFG-NMR measurements. Thus, we reasoned that this value was consistent with a collapsed tetrahedral cage. The solvodynamic radius of **CC-Octane** is smaller than the estimated size of a [4+6] cage, confirming that it most likely adopts a [2+3] topology.

In the series of "odd" cages, the value of ρ_r was found to increase dramatically between three and seven methylene carbons, which we rationalise is an effect of increasing anisotropy in molecular structure. This has been known to increase the observed size of molecules in solution (see Section 2.1.4). By comparison, **CC-nonane** was observed to have a more moderate ρ_r value of 0.85. We rationalise that a structurally-persistent prismatic cage would generate a value of ρ_r greater than 2. Conversely, a fully collapsed cage might be expected to generate a significantly smaller ρ_r value as it adopts a more condensed structure (Figure 2.36). Thus, we rationalise that this value represents an equilibrium mixture between collapsed and inflated conformers.

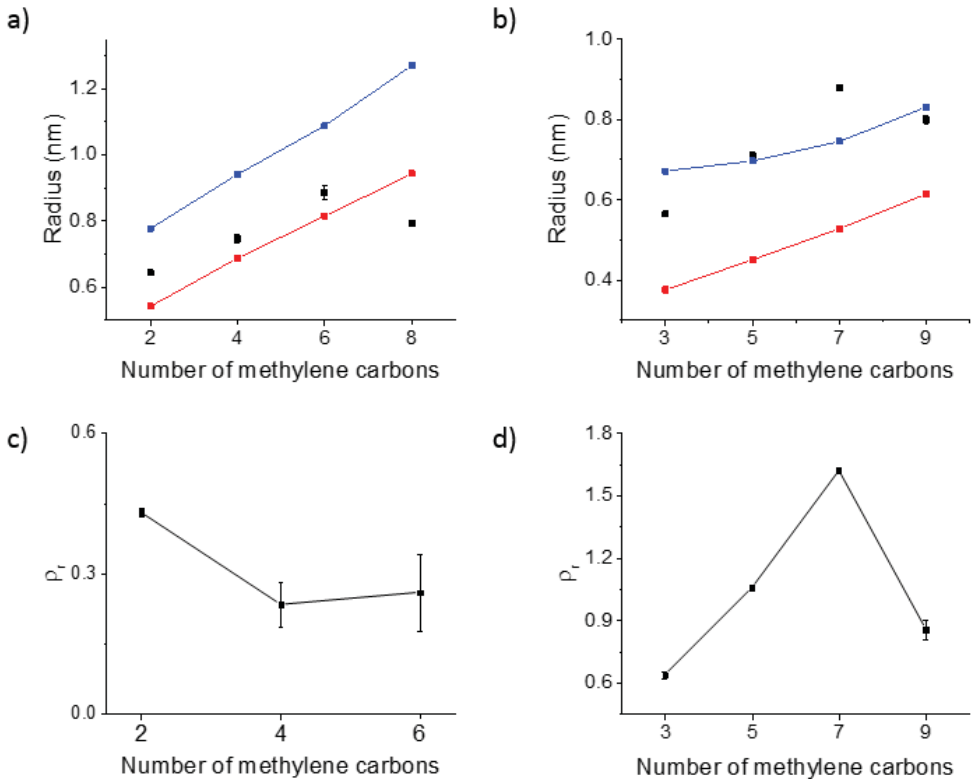


Figure 2.36. PFG-NMR data (black squares) for (a) "even" and (b) "odd" cages compared to r_{acg} (red squares) and r_{max} (blue squares); Values of ρ_r for (c) "even" and (d) "odd" cages.

2.2.2 PFG NMR of cage molecules: a high-throughput tool

To apply ^{1}H PFG-NMR spectroscopic analysis to a wide range of syntheses, and test the analytical capabilities of this technique, it was deployed as part of a collaborative project, which hoped to screen a broad range of monomers with the hope of accelerating cage discovery. This project involved a large number of people and techniques, of which PFG-NMR was only a small part (Figure 2.37). All synthesis here was carried out by Becky Greenaway (UoL), and ^{1}H PFG-NMR spectroscopic analysis was performed on crude reaction mixtures.

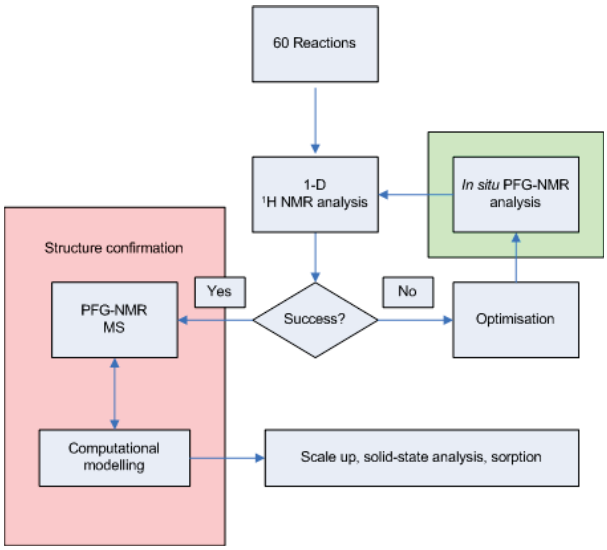


Figure 2.37. High-throughput screening of cage precursors enabled accelerated discovery in collaboration with computational and crystallographic specialists. Highlighted in red is the comparison between computational and PFG-NMR models that facilitates structure characterisation. Highlighted in green is PFG-NMR analysis used to optimise reactions.

Most importantly, ^{1}H PFG-NMR spectroscopic analysis was used as a complementary tool, in hand with high-resolution mass spectrometry (HRMS) in establishing whether reaction mixtures had generated covalent cages and to categorise these cages by topology.

2.2.2.1 Generation of a library of molecules

In generating the library of data used for spectroscopic analysis, 21 aldehydes with 2 or 3 functional moieties were reacted with a small family of amines (Figure 2.38). Aldehyde precursors for reactions are shown in Table 2.4 and Table 2.5.

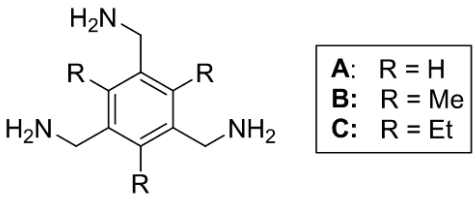


Figure 2.38. Chemical structure of amines A–C used in the high-throughput screen.

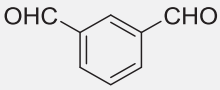
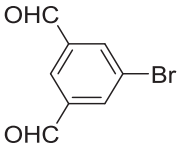
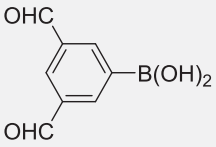
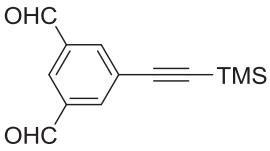
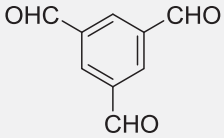
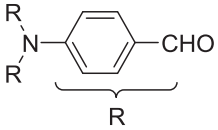
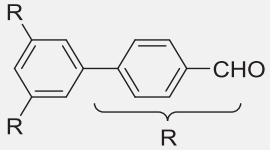
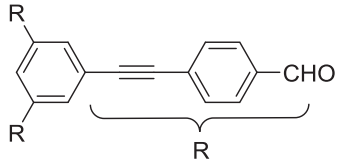
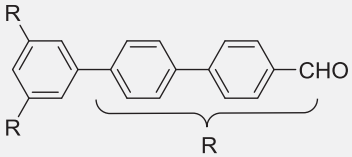
	Aldehyde	A	B	C
1		$A1_{[3+2]}$	$B1_{[3+2]}$	$C1_{[3+2]}$
2		$A2_{[3+2]}$	$B2_{[3+2]}$	$C2_{[3+2]}$
3		$A3_{[3+2]}$	$B3_{[3+2]}$	$C3_{[3+2]}$
4		$A4_{[3+2]}$	$B4_{[3+2]}$	$C4_{[3+2]}$
5		$A5_{[4+4]}$	$B5_{[4+4]}$	$C5_{[4+4]}$
6		$A6_{[4+4]}$	$B6_{[4+4]}$	$C6_{[4+4]}$
7		$A7_{[4+4]}$	$B7_{[4+4]}$	$C7_{[4+4]}$
8		$A8_{[4+4]}$	$B8_{[4+4]}$	$C8_{[4+4]}$
9		$A9_{[4+4]}$	$B9_{[4+4]}$	$C9_{[4+4]}$

Table 2.4. Precursors used in high-throughput screening together with naming scheme for [3+2] and [4+4] cage products.

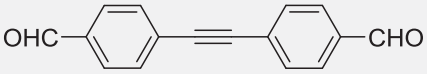
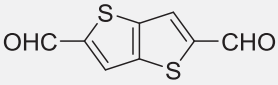
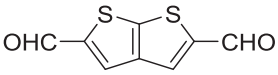
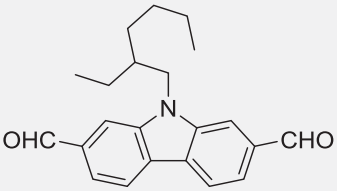
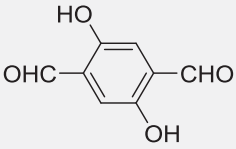
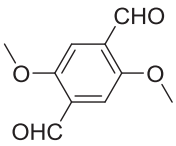
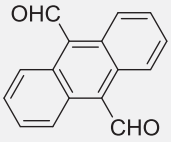
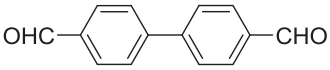
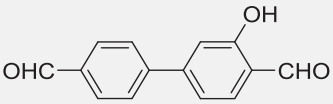
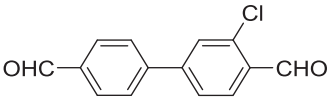
	Aldehyde	A	B	C
10		A10$_{[6+4]}$	B10$_{[6+4]}$	C10$_{[6+4]}$
11		A11$_{[6+4]}$	B11$_{[6+4]}$	C11$_{[6+4]}$
12		A12$_{[6+4]}$	B12$_{[6+4]}$	C12$_{[6+4]}$
13		A13$_{[6+4]}$	B13$_{[6+4]}$	C13$_{[6+4]}$
14		A14$_{[6+4]}$	B14$_{[6+4]}$	C14$_{[6+4]}$
15		A15$_{[6+4]}$	B15$_{[6+4]}$	C15$_{[6+4]}$
16		A16$_{[6+4]}$	B16$_{[6+4]}$	C16$_{[6+4]}$
17		A17$_{[6+4]}$	B17$_{[6+4]}$	C17$_{[6+4]}$
18		A18$_{[6+4]}$	B18$_{[6+4]}$	C18$_{[6+4]}$
19		A19$_{[6+4]}$	B19$_{[6+4]}$	C19$_{[6+4]}$
20		A20$_{[6+4]}$	B20$_{[6+4]}$	C20$_{[6+4]}$
21		A21$_{[6+4]}$	B21$_{[6+4]}$	C21$_{[6+4]}$

Table 2.5. Precursors used in high-throughput screening together with naming schema for [6+4] cage products.

Syntheses were conducted by Becky Greenaway, all using $CDCl_3$ as solvent, which facilitated fast ^1H NMR spectroscopic analysis without the need for solvent suppression. 1D ^1H NMR spectra were acquired on crude samples and analysed prior to PFG-NMR analysis.

As has been previously discussed (see Section 2.1.2), a moderate s/n is required for PFG-NMR analysis. To this end, several factors were considered in the initial 1D ^1H NMR screen in determining whether it was possible to generate statistically significant results through analysis of the PFG-NMR data.

Differences exist in the number of ^1H NMR peaks associated with cages, However, cage symmetry combined with structural flexibility results in small numbers of chemical shifts. In the case of **CC3** a single chemical shift is associated with *H*C=N imine protons. The additional complexity associated with the [8+12] structure of **CC7** generates two imine environments (Figure 2.39).

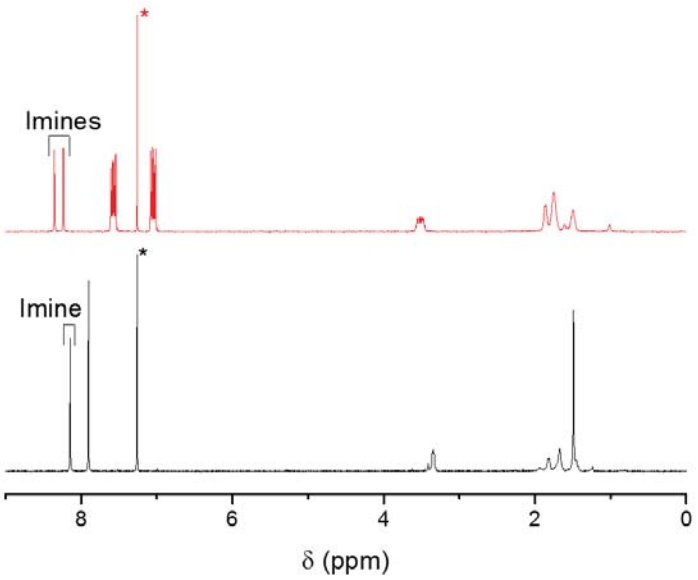


Figure 2.39. ^1H NMR spectra of CC7 (red) and CC3 (black). Solvent peaks associated with $CDCl_3$ have been truncated at half-height for clarity. In each case, this is marked with an asterisk.

As analysis was performed on crude mixtures, chemical shifts were predicted to overlap with starting materials, intermediates and oligomeric species. In cases where large numbers of chemical shifts were observed in the ^{1}H NMR spectrum, peak breadth was an important factor in determining whether PFG-NMR spectroscopy would afford reliable results (Figure 2.40).

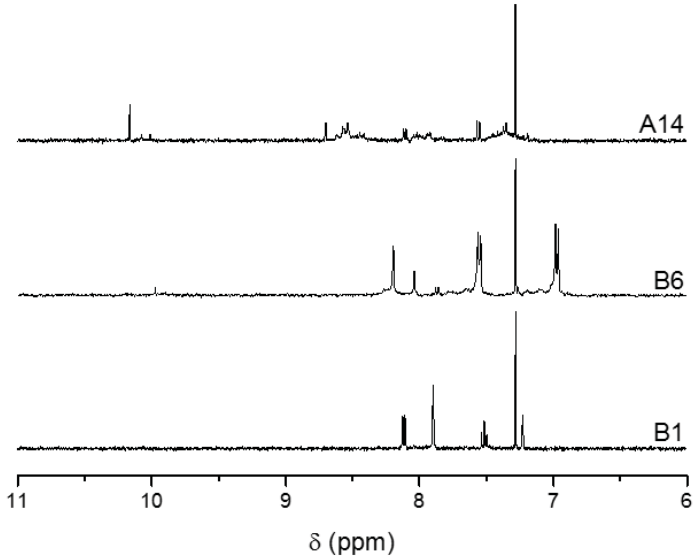


Figure 2.40. ^{1}H NMR spectra of crude reaction mixtures in the attempted synthesis of A14 (top), B6 (middle) and B1 (bottom).

Figure 2.40 demonstrates three "grades" of crude ^{1}H NMR spectra, focussed around the aromatic area of the spectrum. **B1**, which appeared to generate a small number of chemical shifts appeared to indicate a single product. **B6** demonstrated peak broadness, which may indicate additional products or intermediates. **A14** contained chemical shifts at a large range of δ, and represented the lowest s/n that could be reproducibly analysed by PFG-NMR. The quality of each spectrum was determined by observing the noise associated with PFG-NMR response curve, where high s/n is indicated by a straight-line decay, such as in the example of **B1** (Figure 2.41a). Lower s/n is characterised by random deviation from this decay. For example, in the case of

A14, significant noise was noted in points above g^2 = 500 m^2 cm^{-2} (Figure 2.41b). This indicated that values below 500 m^2 cm^{-2} were a reliable indicator of cage size. PFG-NMR response curves demonstrating noise greater than this were not considered statistically significant.

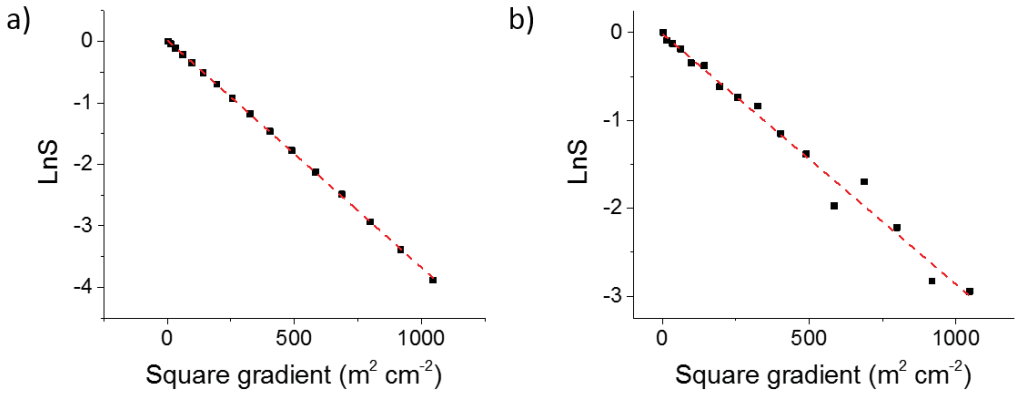


Figure 2.41. PFG-NMR response curves for crude reaction mixtures (a) A

2.2.2.2 Analysis of cage species

A14 demonstrated several downfield chemical shifts at ~ 10 ppm, which were associated with aldehyde protons, and indicated a number of molecules with unreacted aldehyde moieties.

PFG-NMR of the chemical shift at 10.17 ppm confirmed the presence of a small structure, with a solvodynamic radius of 0.6 nm. The s/n associated with chemical shifts upfield of this region (7 – 9 ppm) was lower (s/n = 1.6, Figure 2.40), but indicated species with a solvodynamic radius of ~ 1.3 nm (Figure 2.42), consistent with oligomeric intermediates approximately the size of a small cage.

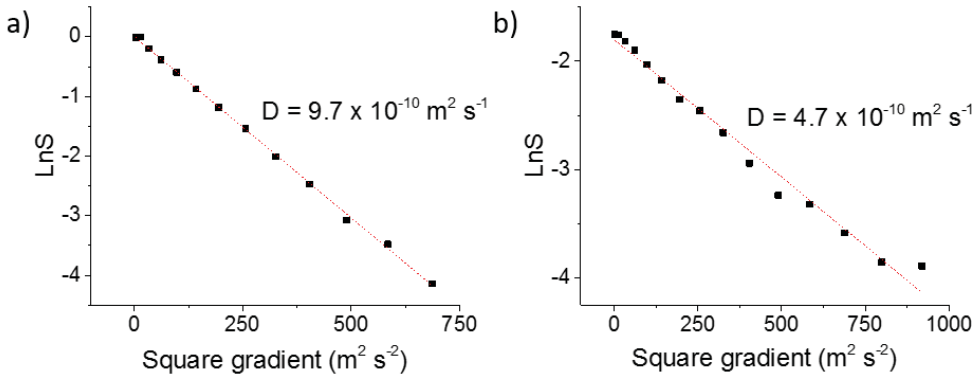


Figure 2.42. PFG-NMR response curves for chemical shifts at (a) 10.17 and (b) 10.01 ppm in the crude ^{1}H NMR spectrum obtained during the synthesis of A14.

This difference in s/n is observed on treating diffusion data with the Inverse Laplace transformation (ILT) required to produce a 2D DOSY spectrum. Chemical shifts in the region 10.1 – 10.0 ppm do not appear in the DOSY spectrum, as results in this region do not have statistical significance, resulting in broad DOSY peaks (Figure 2.43).

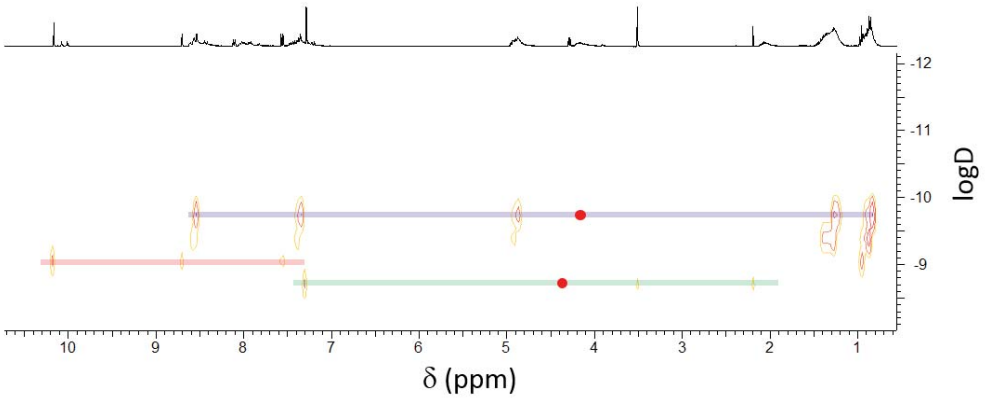


Figure 2.43. ^{1}H DOSY spectrum of crude reaction mixture obtained in the synthesis of A14. Chemical shifts which share diffusion coefficients have been highlighted. Red circles indicate DOSY peaks too broad to appear at this resolution due to ^{1}H NMR chemical shift overlap.

In spite of this, the analysis can confirm which chemical shifts are the product of single molecular species. Chemical shifts at 10.17, 8.72 and 7.58 ppm appear to be generated by a single oligomeric species of R_S = 0.6 nm (highlighted in red, Figure 2.43).

Chemical shifts at 8.55, 7.38, 4.90, 1.23 and 0.85 appeared to be generated by a second, larger species (highlighted in blue, Figure 2.43), which was 2.1 nm in diameter. This is consistent with a macromolecular cage. Furthermore, the presence of chemical shifts at ~ 8 ppm indicated the possible presence of imine protons. No aldehyde environments (those at ~10 ppm) were observed at this diffusion coefficient, which indicated that all aldehydes in the structure had reacted to form imines.

To confirm the geometry and characterise the structure of **A14**, diffusion data was compared with the predicted structure produced by Valentina Santolini at ICL (Figure 2.37). Computational predictions suggested a [3+2] cage, with r_{max} = 1.41 nm, and r_{avg} = 0.71 nm, would be formed by the reaction of amine **14** and aldehyde **A**. This structure was consistent with the size of the species being measured by diffusion NMR. R_S was found to fall mid-way between r_{max} and r_{avg}, with ρ = 0.48, which indicated an incurvate structure (see Section 2.2.1.2). Thus, from a mixed spectrum, PFG-NMR facilitated ^{1}H NMR spectrum assignment (Figure 2.44).

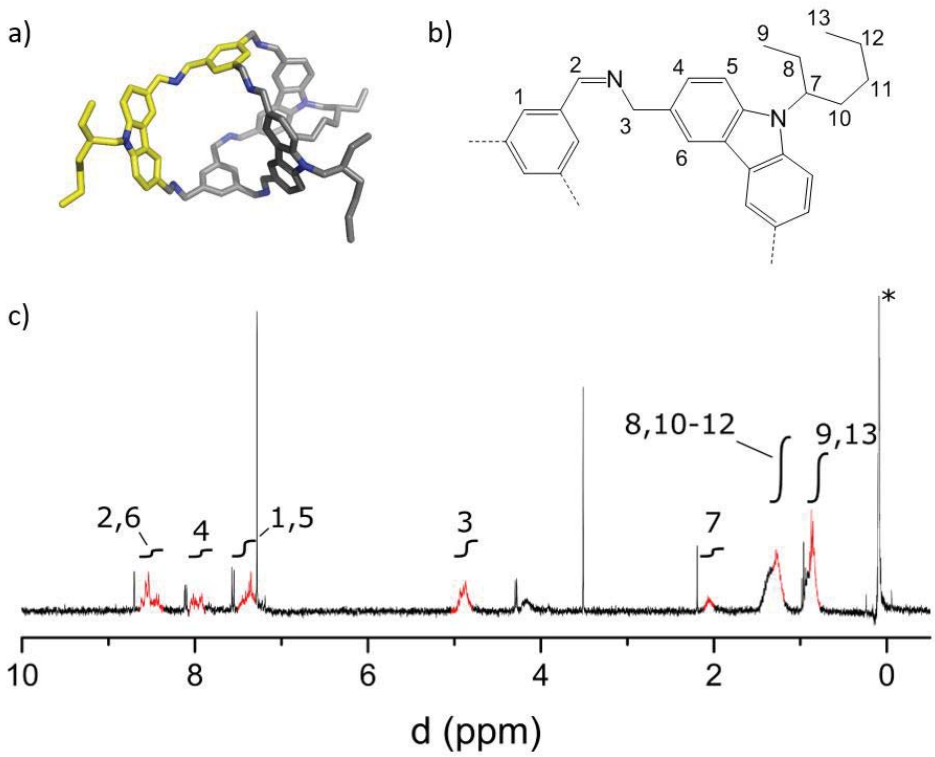


Figure 2.44. (a) Computationally-minimised structure of A14. Highlighted in yellow is the section of the chemical structure that corresponds to (b); (c) ^{1}H NMR spectrum of A14. Numbers correspond to chemical shift assignment.

Using this above approach, crude ^{1}H NMR spectra were screened jointly with Becky Greenaway, separating crude reaction mixtures that could be analysed by PFG-NMR from those that could not. In this way, PFG-NMR was performed on 27 of 60 crude reactions.

2.2.3 High-throughput evaluation of cage size

Samples that contained one macromolecular species by ^{1}H NMR and those that contained mixtures of products, but where chemical shifts were resolved, were analysed by PFG-NMR. Where s/n was deteriorated by precipitation of reaction

species, or complex mixtures (grey, Table 2.6), PFG-NMR was not conducted on samples.

^{1}H NMR screen					
A1	B1	C1	A11	B11	C11
A2	B2	C2	A12	B12	C12
A3	B3	C3	A13	B13	C13
A4	B4	C4	A14	B14	C14
A5	B5	C5	A15	B15	C15
A6	B6	C6	A16	B16	C16
A7	B7	C7	A17	B17	C17
A8	B8	C8	A18	B18	C18
A9	B9	C9	A19	B19	C19
A10	B10	C10	A20	B20	C20
			A21	B21	C21

Table 2.6. Success of reactions, classified by (green) a single product, (yellow) a mixture of products, one of which was consistent with cage material and (grey) mixtures too complex or weak to facilitate analysis.

Subsequently, PFG-NMR data was compared to molecular models generated by Valentina Santolini (ICL). Three cage topologies were studied: [3+2], [6+4] and [4+4]. Cages that were expected to fall into specific topologies have been grouped, and are dealt with together. Several examples are discussed in which the predicted topology did not agree with the experimental results.

2.2.3.1 [3+2] cage series

Of reactions that were predicted to form [3+2] cages, six were found to produce ^{1}H NMR spectra that indicated the presence of imine groups, and were of sufficient strength to generate meaningful PFG-NMR data (Figure 2.45).

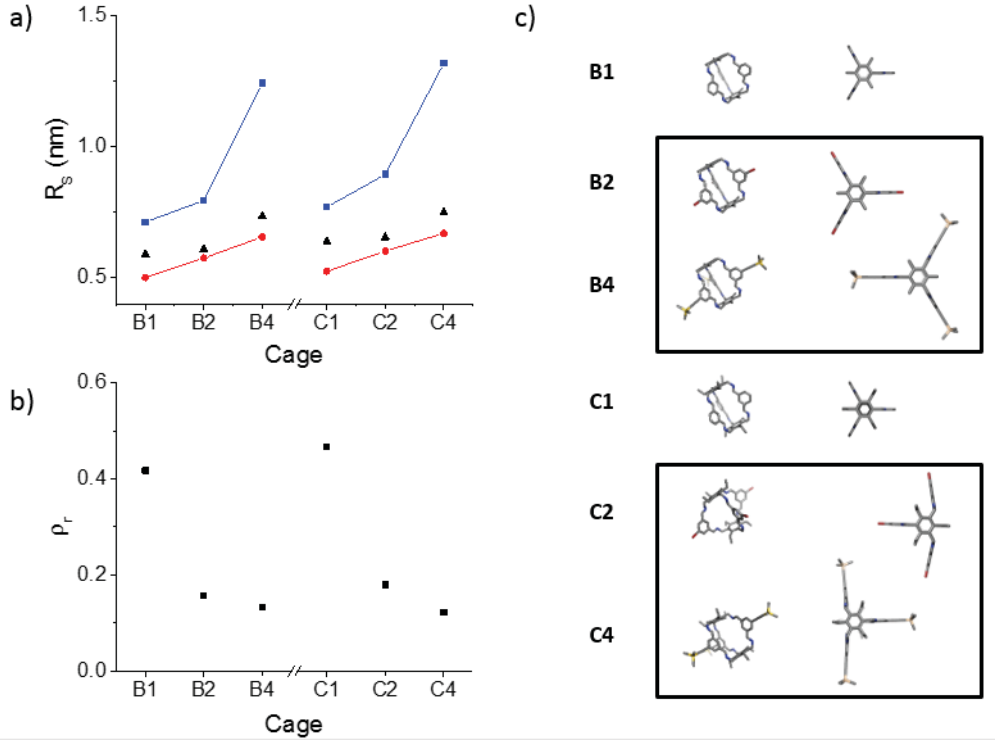


Figure 2.45. (a) Solvodynamic radii (black triangles) of compounds expected to form [3+2] supramolecular cages. Values of r_{max} (blue squares) and r_{avg} (red circles) were calculated from computationally-generated structures, which are shown in (c). Error bars are smaller than the size of data point. (b) Values of ρ_r. (c) Computationally-generated structures of cages, shown side-on (left) and down the C_3 axis (right).

PFG-NMR data was found to agree with computationally-generated structures, lying between r_{avg} and r_{max} in all six experiments. The solvodynamic radii of cages synthesised from triamine **C** were observed to be on average 37 pm larger than those synthesised from triamine **B**, but with the corresponding aldehyde. This was attributed to the presence of the ethyl groups in triamine **C**. **B1** and **C1**, which typify the

short-prism geometry of **CC-pentane** and **CC-propane**, were found to have ρ_r values of 0.42 and 0.47. These we compare to the ρ_r value of 0.6 obtained for **CC-propane**. In explaining the difference, we noted that these prismatic structures have, on average, a longer axial dimension than an equatorial one, which is extended in one dimension where **CC-propane** was extended in two.

Similarly, **B2**, **B4**, **C2** and **C4** were all found to have ρ_r values ranging from 0.12 to 0.18. This was attributed to the functional groups at the *meta*-position on the phenyl rings of the aromatic aldehyde starting materials. We reason that these "winged" prisms have low ρ_r values because, unlike in the case of **Astetrapod** where the "apexes" widen, these projections are narrow, which increases the size of r_{max} without affecting the size of r_{avg}.

2.2.3.2 [4+4] cage series

Furthermore, of reactions that were predicted to form [4+4] cages, eight were found to produce ^{1}H NMR spectra that indicated the presence of imine groups, and were of sufficient strength to generate meaningful PFG-NMR data (Figure 2.46).

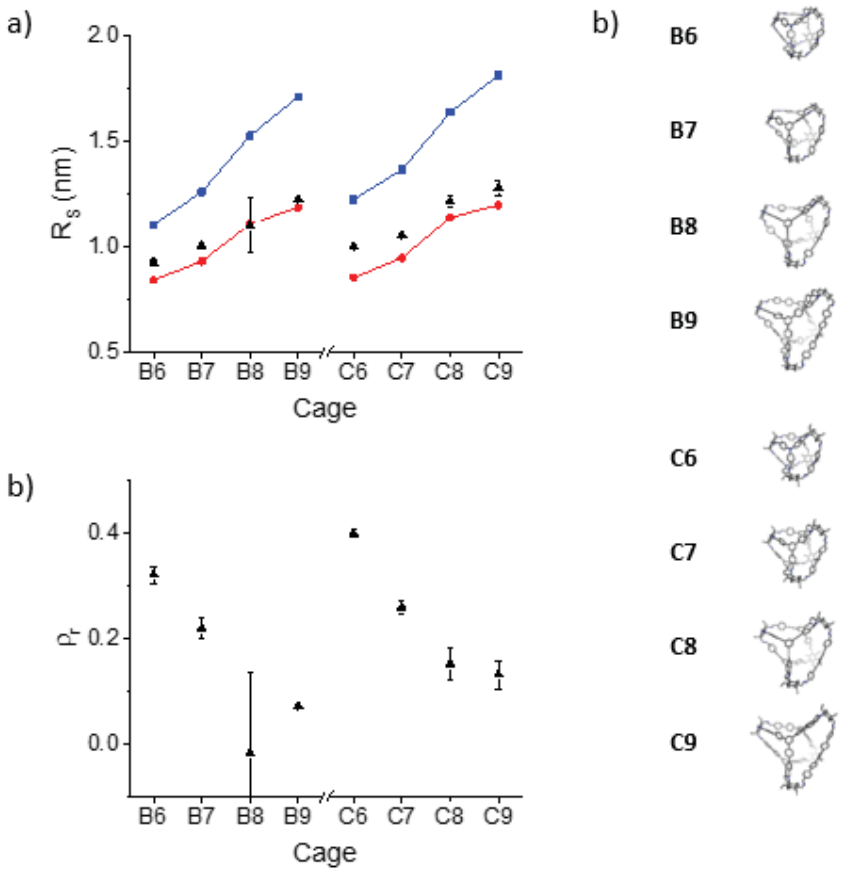


Figure 2.46. (a) Solvodynamic radii (black triangles) of compounds expected to form [4+4] supramolecular cages. Values of r_{max} (blue squares) and r_{avg} (red circles) were calculated from computationally-generated structures, which are shown in (c). (b) Values of ρ_r. (c) Computationally-generated structures of cages, shown side-on using a capped-sticks representation.

Again, PFG-NMR data were found to agree with r_{max} and r_{avg} calculated from computationally-generated structures (Figure 2.46a). We reasoned that the presence of ethyl groups on the central phenyl ring, which were at the end of the four "apexes" in the [6+4] topology should have approximately twice the effect on R_s when compared to the short prisms discussed in Section 2.2.2.1. Correspondingly, cages **C6–9** synthesised from triamine **C** were observed to be ~ 60 pm larger than cages that were synthesised from triamine **B** with the corresponding aldehyde. Values of ρ_r were

found to decrease as the size of the aldehyde constituent was increased from **B6** to **B9**, and **C6** to **C9** (Figure 2.46b). This indicated a significant change in cage shape as the aldehyde was incrementally extended. This corresponded with computational models, which indicated Φ_S decreased from 0.80 to 0.49 as cages became more stellated (Figure 2.47). However, this represented an opposing trend to that noted in Section 2.2.1.1 (Figure 2.20). This was not examined, but we reasoned that this could indicate an increase in the flexibility of larger cages, which would result in a smaller-than-expected R_S, and could explain the decrease in ρ_r. Structural flexibility was also expected to reduce the rigidity of "apexes" in the molecule, which would increase Φ_S in larger species.

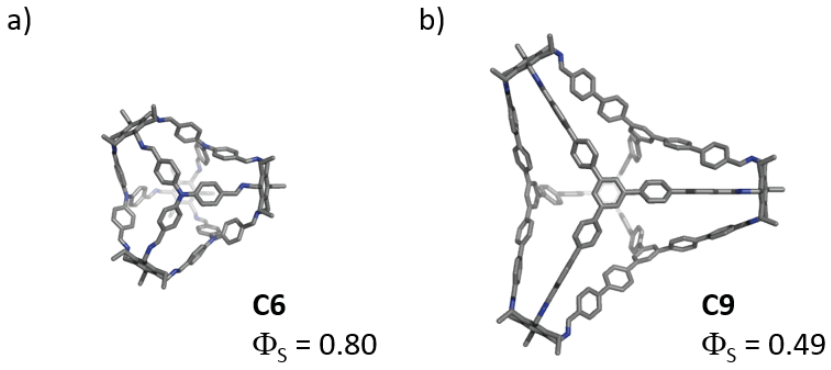


Figure 2.47. Molecular structure of (a) C6 and (b) C9 showed as capped-sticks representations, with hydrogen atoms excluded for clarity.

2.2.3.3 [6+4] cage series

Finally, 13 reactions were identified in which the products were expected to form cages adopting a [6+4] topology, and in which the ^{1}H NMR spectra indicated either a single product or major component was produced. Unlike in the case of [3+2] and

[4+4] cages, experimental results were observed to provide a much poorer fit with computationally-generated structures, and ρ_r was observed to vary considerably.

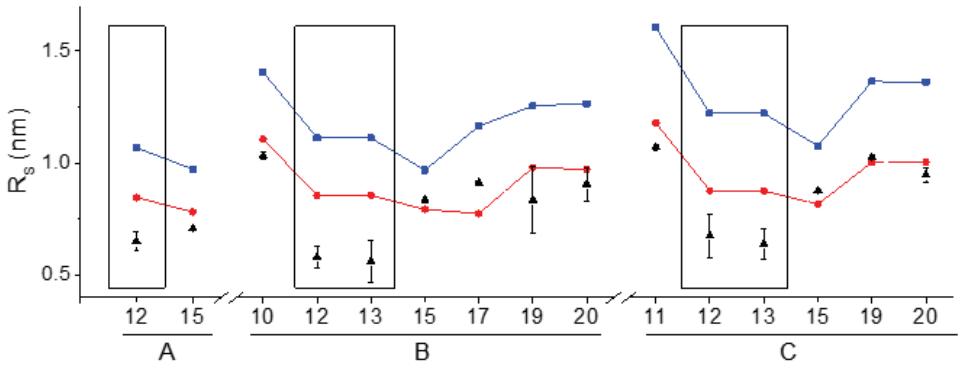


Figure 2.48. (a) Solvodynamic radii (black triangles) of compounds expected to form [6+4] supramolecular cages. Values of r_{max} (blue squares) and r_{avg} (red circles) were calculated from computationally-generated structures, which are shown in (b). In some cases, the error associated with r_s is smaller than the size of the data point used. Three anomalous results, in which r_s is significantly smaller than expected, have been highlighted using black boxes.

R_S calculated from PFG-NMR analysis of **A12, B12, C12, A15**, **B19-20** and **C20**, were less than r_{avg}. In the case of reactions **A–C12 and B–C13**, this difference was significant, indicating that PFG-NMR analysis was not consistent with a cage of [6+4] topology. Mass spectra collected by Becky Greenaway did not demonstrate peaks consistent with either oligomeric materials or covalent cages, although mass ions in the region of 800 *m/z* indicated reaction products with a smaller number of components than expected. On the basis of their smaller measured size, we reasoned that **B13** and **C13** may have formed prismatic [3+2] cages, which would have a lower molecular mass. To test this hypothesis, molecular models of **B12** and **C12** were generated for the [3+2] topology (Figure 2.49) by Valentina Santolini (ICL). Solvodynamic radii were found to lie between r_{max} and r_{avg} calculated from the structures, and ρ_r was found

to be consistent with a short-prism topology. This could be investigated further by comparing diffusion data to computational models for **A–C13**. However, at the time of writing, these had not been produced. Similarly, cages **B19–20** and **C19–20** were identified as cages with [6+4] topology in spite of smaller-than-expected values of r_s.

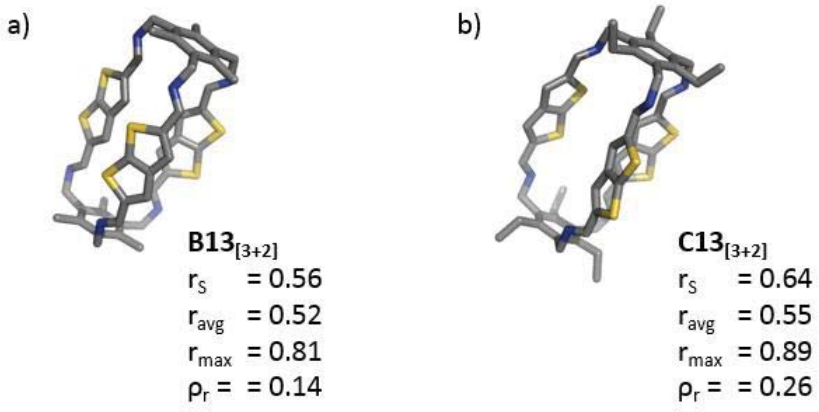


Figure 2.49. Computationally-generated structures of (a) B13 and (b) C13 shown as a capped-sticks representation. Values of r_x, r_{avg}, r_{max} and ρ_r are shown for each structure. Hydrogen atoms have been omitted for clarity.

The ^{1}H NMR spectrum of **A15** demonstrated several clusters of chemical shifts with complex multiplicity. These were determined by diffusion NMR to be the result of smaller species in solution, having a solvodynamic radius of 0.4 nm. Peak overlap with chemical shifts at 8.39, 7.71, 4.89 and 3.91 ppm that were associated with cage **A15** (Figure 2.50a) resulted in a non-linear PFG-NMR response curve, resulting from the linear combination of two response curves at the same chemical shift (Figure 2.50b). Attempts to fit this to a single diffusion coefficient resulted in smaller-than-expected solvodynamic radii. To avoid this, we fitted the data for **A15** to a two-component PFG-NMR response curve, which generated two solvodynamic radii of 0.42 nm and 0.78nm (Figure 2.49c). We reasoned that this may still represent an

underestimation of r_s, as it was unknown whether the "minor component" was composed of a single molecular structure, or several oligomeric species. Addition of subsequent response curves can complicate resolution, which is discussed in Chapter 3, Section 3.1.1. On this basis, cages **A15**, **B15** and **C15** were assigned as [6+4] cages.

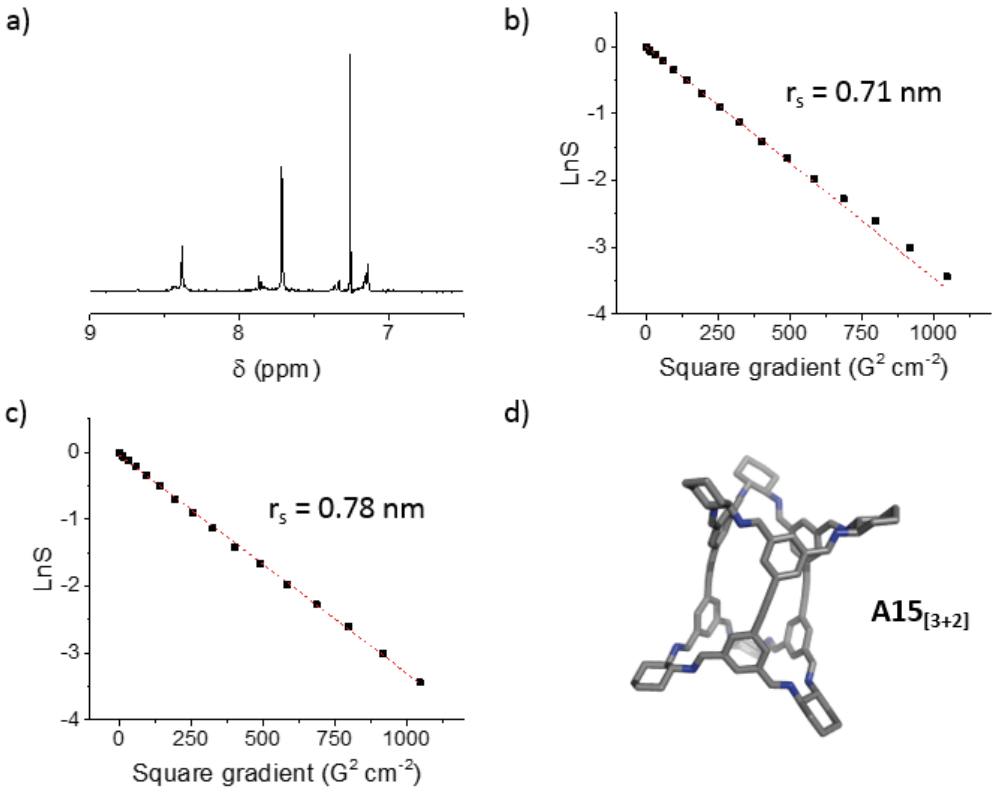


Figure 2.50. (a) Crude ^{1}H NMR spectrum of A15; PFG-NMR response curves fitted using (a) a straight line and (b) a two-component response curbe; (d) Molecular structure of A15 shown as a capped sticks representation. Hydrogen atoms have been omitted for clarity.

In conclusion, 27 cages were identified and characterised by PFG-NMR analysis in a high-throughput screen of 20 aldehydes and 3 triamines. Cages that adopted a prismatic geometry demonstrated good agreement with the expected values of ρ_r, whilst tetrapodal [4+4] cages demonstrated a p_r values that deviated from anything previously measured. This was reasoned to be the result of increasing cage flexibility,

which could be studied both computationally, and by varying the solvents in which the PFG-NMR was conducted (see Section 2.2.1). Tetrahedral cages were more complex to analyse and in several cases diffusion NMR results indicated that cages may adopt a topology different to that which had been predicted. In future work, it would be useful to confirm these findings by isolating material, rather than conducting analysis on crude reaction mixtures. Solvent probe experiments (see Section 2.2.1.2) could be used to characterise the curvature of the molecular surface, which could also be used to identify the likely topology.

2.3 Conclusions

In this Chapter, two techniques have been developed, which allow structural details of macromolecular cages to be determined in solution. The degree of curvature associated with the solvent-accessible surface of a molecule can be probed, which can afford detailed structural insights about the topology of a cage. Prismatic cages, for instance, are convex, and changes in solvent are unlikely to result in noticeable changes to D.

Two series of molecules have been characterised by diffusion NMR. The degree of structural collapse was estimated by comparing the diffusion coefficients of similar molecules, as well as generating PFG-NMR data from computational models. This technique was incorporated into a high-throughput screen, which facilitated the analysis and characterisation of 27 new cage molecules.

2.4 Experimental

All reagents obtained from Acros, Aldrich, TCI Europe, and Manchester Organics fine chemicals suppliers were used directly as supplied, or dried by the methods described by Burfield *et al.*[90] Where materials were dried, this is identified in the

experimental of that compound. Unless otherwise indicated, temperatures included for heating and cooling were measured externally.

NMR Spectra: ^{1}H NMR Spectra were recorded using an internal deuterium lock for the residual protons in $CDCl_3$ (δ 7.26), methanol-*d*$_4$ (3.34 ppm) or CD_2Cl_2 (5.50 ppm) at 298 K on the following instruments: Bruker AV400 (400 MHz), Bruker Avance 400 DPX (400 MHz) or Bruker DRX500 (500 MHz). Data are presented as follows: chemical shift, integration, peak multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, qu = quintet, m = multiplet, br = broad), coupling constants (*J* / Hz), and assignment. Chemical shifts are expressed in ppm on a δ scale relative to δ_{TMS} (δ = 0 ppm) and coupling constants, *J*, are given in Hz. Assignments were determined either on the basis of unambiguous chemical shift, by analysis of 2D NMR (COSY, HSQC, HMBC), or by analogy to fully interpreted spectra for structurally related compounds. ^{13}C NMR Spectra were recorded using an internal deuterium lock using $CDCl_3$ (δ 77.0) at ambient probe temperatures on the following instruments: Bruker AV400 (101 MHz), Bruker Avance 400 DPX (101 MHz) or Bruker DRX500 (126 MHz).

PFG-NMR spectra: All measurements were carried out non-spinning on a 400 MHz Bruker Avance 400 spectrometer, using a 5 mm indirect detection probe, equipped with a z-gradient coil producing a nominal maximum gradient of 34 G cm^{-1}. Diffusion data was most frequently collected using the Bruker longitudinal eddy current delay (LED) pulse sequence (ledgp2s).[91] In the case of highly concentrated, viscous samples, bipolar gradients were used (ledbpgp2s)[92] to minimise artefacts in the spectrum. In the case of highly convective samples, the Bruker double stimulated-emission pulse sequence (dstegp3s)[59,60] was used to remove the effects of convection.

A diffusion encoding pulse δ of length 1–7 ms, and diffusion delay Δ of 0.1–0.35 s were used. Gradient amplitudes were spaced between 1.70 and 32.4 G cm^{-1}. In cases for simple diffusion measurements, these were spaced linearly. Where dispersity was estimated using this technique, gradient amplitudes were spaced exponentially. Diffusion NMR measurements were obtained using a double-stimulated emission pulse sequence (dstegp3s)[59,60] to remove the effects of convection.

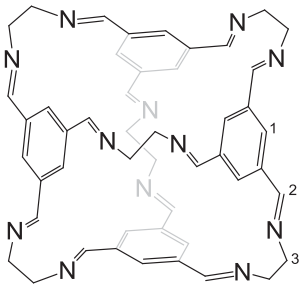
Each FID was acquired using 16 k data points. All experiments were carried out at a nominal probe temperature of 298 K, with an air flow of 800 m^3 min^{-1} to minimise sample heating.

Mass Spectra: Low resolution mass spectra (*m/z*) were recorded on a Micromass LCT Mass Spectrometer or a Trio 1000 Quadrupole Mass Spectrometer. High resolution mass spectra (*m/z*) were recorded on an Agilent Technologies 6530B accurate-mass QTOF Dual ESI mass spectrometer in positive-ion detection mode. The mobile phase was MeOH with 0.1% formic acid at a flow rate of 0.25 mL min^{-1}.

Other techniques: Reactions were monitored by thin layer chromatography (TLC) or by high performance liquid chromatography (HPLC). TLC was conducted on pre-coated aluminium-backed plates (Merck silica gel 60 with fluorescent indicator UV254). Spots were visualized either by quenching of UV fluorescence (254 nm) or by staining with potassium permanganate dip. HPLC was conducted on a Dionex UltiMate 3000. Flash column chromatography was performed either manually according to the method described by Still, Khan and Mitra[93] with silica gel 60 (40-63 μm), applying head pressure by means of nitrogen, or on a Biotage Isolera Four with Biotage Snap KP-Sil normal phase disposable columns.

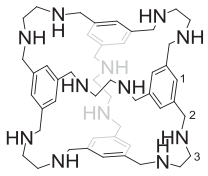
2.4.1 Synthesis of cages

Synthesis of CC1:



To a cooled (0 °C) solution of benzene-1,3,5-tricarbaldehyde (450 mg, 2.78 mmol, 1.0 eq.) in DCM (100 mL) was added dropwise a solution of ethane-1,2-diamine (250 mg, 4.16 mmol, 1.5 eq.) in DCM (130 mL) over 18 hours. The resultant clear solution was allowed to warm to r.t. and maintained at r.t. for 24 hours to afford a clear, colourless solution. Solvents were removed *in vacuo* to afford a fine white powder (677 mg, 2.70 mmol, 97%). **MS (ESI)** m/z = 793.4 $[M+H]^+$; **^{1}H NMR** (400 MHz, $CDCl_3$): δ ppm 8.19 (12H, s, H2), 7.93 (12H, s, H1), 4.03 (24H, s, H3). Data in accordance with literature values.[40]

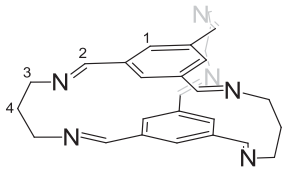
Synthesis of RCC1:



To a stirring solution of **CC1** (500 mg, 0.631 mmol, 1.0 eq.) in MeOH (150 mL) was added portion-wise $NaBH_4$ (770 mg, 20.3 mmol, 32 eq.) and the resultant clear solution was maintained at r.t. for 16 hours. To the solution was added water (2 mL) and the solution was stirred for a further 2 hours. Solvents were removed *in vacuo* to afford an off-white solid. Crude solid was extracted with $CHCl_3$ (200 mL), and the remaining solids were removed by filtration. Organic solvents were removed *in vacuo* to afford a white solid (250 mg, 0.306 mmol, 48%). **MS (ESI)** *m/z* = 817.6 $[M+H]^+$;

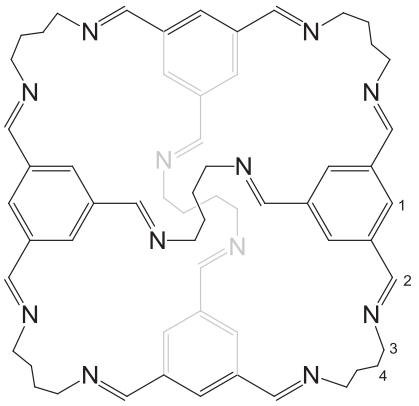
^{1}H NMR ($CDCl_3$, 400MHz) δ ppm 7.07 (12H, s, H1), 3.69 (24H, s, H2), 2.65 (24H, s, H3). Data in accordance with literature values.[41]

Synthesis of CC-propane:



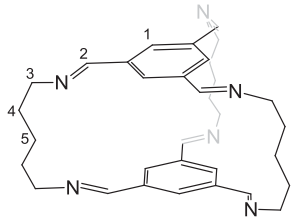
Propane-1,3-diamine was dried over activated 4 Å molecular sieves, in accordance with literature procedures. To a cooled (0 °C) solution of propane-1,3-diamine (54.4 mg, 0.734 mmol, 1.6 eq.) and DMF (1 drop) in DCM (50 mL) was added by syringe pump a solution of benzene-1,3,5-tricarbaldehyde (75.2 mg, 0.464 mmol, 1.0 eq.) in DCM (40 mL) over 24 hours. The resultant mixture was allowed to warm to r.t. and maintained at r.t. for 72 hours to afford a turbid solution. Reaction completion was confirmed by ^{1}H NMR spectroscopic analysis. MS and NMR data were gathered directly from crude reaction mixtures. **MS (ESI)** *m/z* = 439.3 $[M+H]^+$; **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.14 (6H, s, H2) 7.79 (6H, s, H1) 3.89 (12H, t, *J* = 5.0 Hz, H3) 1.24 (6H, qu, *J* = 5.0 Hz, H4). Data in accordance with literature values.[43]

Synthesis of CC-butane:



To a cooled (0 °C) solution of butane-1,4-diamine (64.9 mg, 0.736 mmol, 1.6 eq.) and DMF (1 drop) in DCM (50 mL) was added dropwise a solution of benzene-1,3,5-tricarbaldehyde (76.3 mg, 0.471 mmol, 1.0 eq.) in DCM (40 mL) over 24 hours. The resultant mixture was allowed to warm to r.t. and maintained for 72 hours to afford a turbid solution. Reaction completion was confirmed by ^{1}H NMR spectroscopic analysis. MS and NMR data were gathered directly from crude reaction mixtures. **MS (ESI)** *m/z* = 961.6 $[M+H]^+$; **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.22 (12H, s, H2) 8.01 (12H, s, H1) 3.63 (24H, t, *J* = 6.7 Hz, H3) 1.60 (24H, m, H4). Data in accordance with literature values.[43]

Synthesis of CC-pentane:



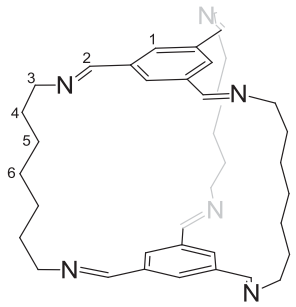
To a cooled (0 °C) solution of pentane-1,5-diamine (425 mg, 4.14 mmol, 1.5 eq.) in DCM (50 mL) was added dropwise a solution of benzene-1,3,5-tricarbaldehyde (450 mg, 2.78 mmol, 1.0 eq.) in DCM (40 mL) over 6 hours. The resultant mixture was allowed to warm to r.t. and maintained at r.t. for 120 hours to afford a clear, light

yellow solution. Solvents were removed *in vacuo* to afford a pale yellow solid (721 mg, 1.38 mmol, quant.). **MS (ESI)** *m/z* = 523.5 ($[M+H]^+$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.09 (6H, s, H2) 7.86 (6H, s, H1) 3.61 (12H, t, *J* = 5.2 Hz, H3) 1.77 (12H, tt, *J* = 8.2, 5.2 Hz, H4) 0.88 (6H, m, H5); **^{13}C NMR** (101 MHz, $CDCl_3$) δ ppm 160.23, 136.24, 128.90, 61.32, 29.72, 28.82, 23.78. Data in accordance with literature values.[43]

Synthesis of CC-hexane:

To a cooled (0 °C) solution of hexane-1,6-diamine (87.0 mg, 0.75 mmol, 1.5 eq.) in DCM (150 mL) was added dropwise a solution of benzene-1,3,5-tricarbaldehyde (81.1 mg, 0.50 mmol, 1.0 eq.) in DCM (50 mL) over 24 hours. The resultant mixture was allowed to warm to r.t. and maintained for 72 hours to afford a turbid solution. The supernatant was separated from precipitate that had formed over the course of the reaction by filtration and reaction completion was confirmed by ^{1}H NMR spectroscopic analysis. MS and NMR data were gathered directly from crude reaction mixtures. **HR-MS (ESI)** $C_{72}H_{97}N_{12}^+$ expt. 1129.7962, calc. 1129.7954. **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.24 (12H, s, N=C*H*Ar) 8.02 (12H, s, Ar*H*) 3.55 (24H, m, NCH_2) 1.65 (24H, m, CH_2) 1.30 (24H, m, CH_2).

Synthesis of CC-heptane:



To a cooled (0 °C) solution of heptane-1,7-diamine (542 mg, 4.16 mmol, 1.5 eq.) in DCM (50 mL) was added dropwise a solution of benzene-1,3,5-tricarbaldehyde (450 mg, 2.78 mmol, 1.0 eq.) in DCM (40 mL) over 6 hours. The resultant mixture was allowed to warm to r.t. and maintained at r.t. for 36 hours to afford a clear, colourless solution. Solvents were removed *in vacuo* to afford an off-white solid, which was dissolved in $CHCl_3$. Solids were removed by filtration and the solvent concentrated *in vacuo* to afford an off-white solid, which was dried *in vacuo* (40 °C) for 72 hours to afford a fluffy white solid (833 mg, 1.37 mmol, 98%). **MS (ESI)** *m/z* = 607.4 ($[M+H]^+$); **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.20 (6H, s, H2) 8.02 (6H, s, H1) 3.60 (12H, t, *J* = 6.0 Hz, H3) 1.68 (12H, tt, *J* = 8.0, 6.0 Hz, H4) 1.28 (12H, tt, *J* = 8.1, 8.0 Hz, H5) 1.05 (6H, qu, *J* = 8.1 Hz, H6).

Synthesis of CC-octane:

To a cooled (0 °C) solution of octane-1,6-diamine (108.2 mg, 0.75 mmol, 1.5 eq.) in DCM (150 mL) was added dropwise a solution of benzene-1,3,5-tricarbaldehyde (81.1 mg, 0.50 mmol, 1.0 eq.) in DCM (50 mL) over 48 hours. The resultant mixture was allowed to warm to r.t. and maintained for 72 hours to afford a turbid solution. The supernatant was separated from precipitate that had formed over the course of the reaction by filtration and reaction completion was confirmed by ^{1}H NMR spectroscopic analysis. MS and NMR data were gathered directly from crude reaction

mixtures. **HR-MS (ESI)** $C_{42}H_{61}N_6^+$ expt. 648.4975, calc. 648.4879. **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.26 (6H, m, N=C*H*Ar) 8.03 (6H, s, Ar*H*) 3.57 (12H, m, NCH_2) 1.67 (12H, br. s, CH_2) 0.85 (24H, m, CH_2).

Synthesis of CC-nonane:

To a cooled (0 °C) solution of nonane-1,7-diamine (79.0 mg, 0.50 mmol, 1.5 eq.) in DCM (50 mL) was added dropwise a solution of benzene-1,3,5-tricarbaldehyde (53.9 mg, 0.33 mmol, 1.0 eq.) in DCM (40 mL) over 6 hours. The resultant mixture was allowed to warm to r.t. and maintained at r.t. for 24 hours to afford a clear, colourless solution. Reaction completion was confirmed by ^{1}H NMR spectroscopic analysis. MS and NMR data were gathered directly from crude reaction mixtures. **HR-MS (ESI)** $C_{45}H_{67}N_6^+$ expt. 691.5422, calc. 691.5438; **^{1}H NMR** (400 MHz, $CDCl_3$) δ ppm 8.26 (6H, s, N=C*H*Ar) 8.10 (6H, s, Ar*H*) 3.62 (24H, t, *J* = 6.47 Hz, NCH_2 + CH_2) 1.68 (24H, m, CH_2) 1.28 (72H, m, CH_2).

2.4.2 Spectra for novel compounds:

Spectra for CC-propane:

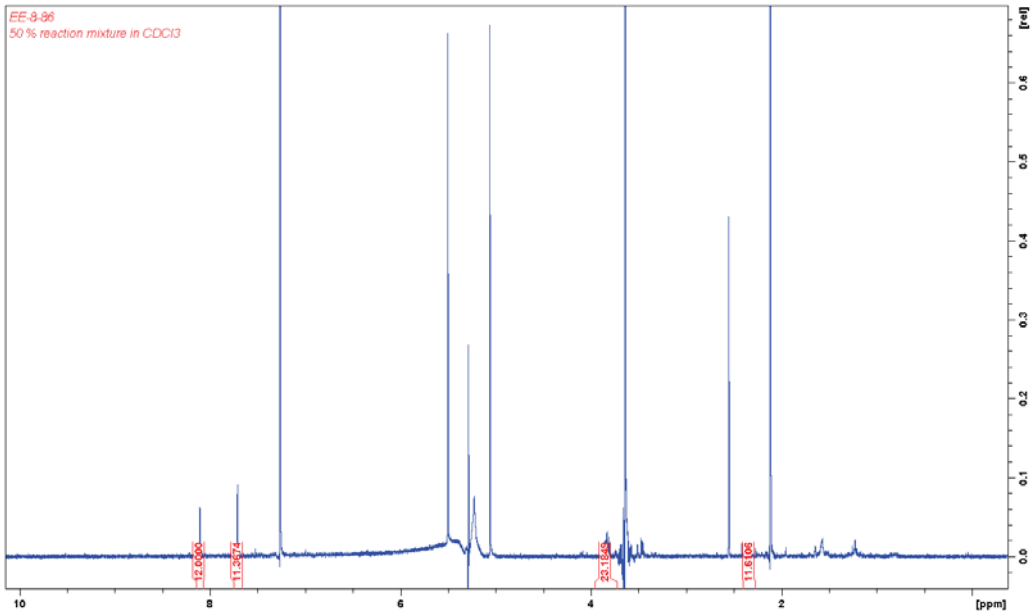


Figure 2.51. ^{1}H NMR spectrum of CC-propane

Spectra for CC-butane:

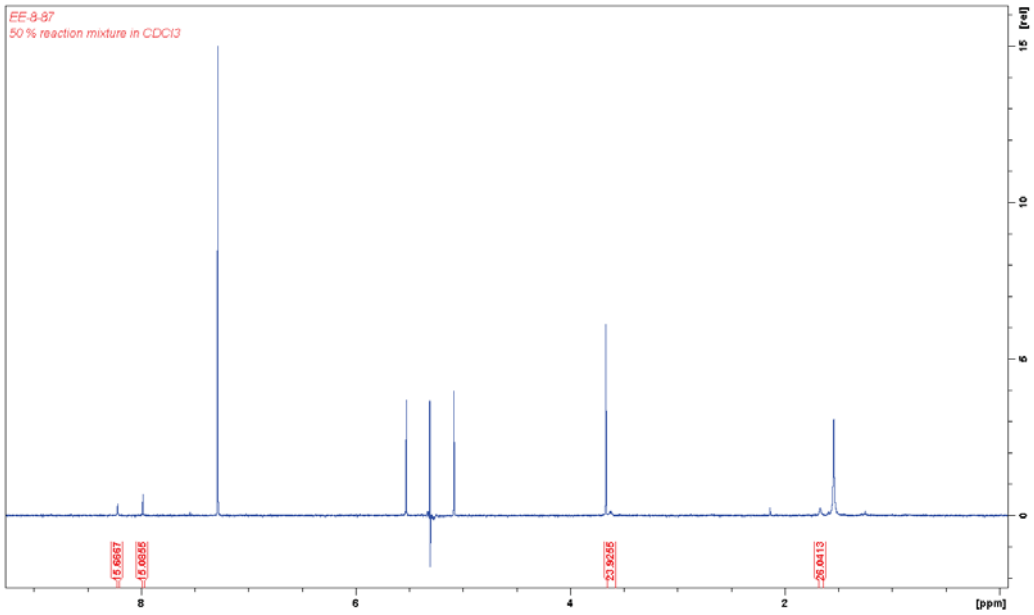


Figure 2.52. ^{1}H NMR spectrum of CC-butane

Spectra for CC-pentane:

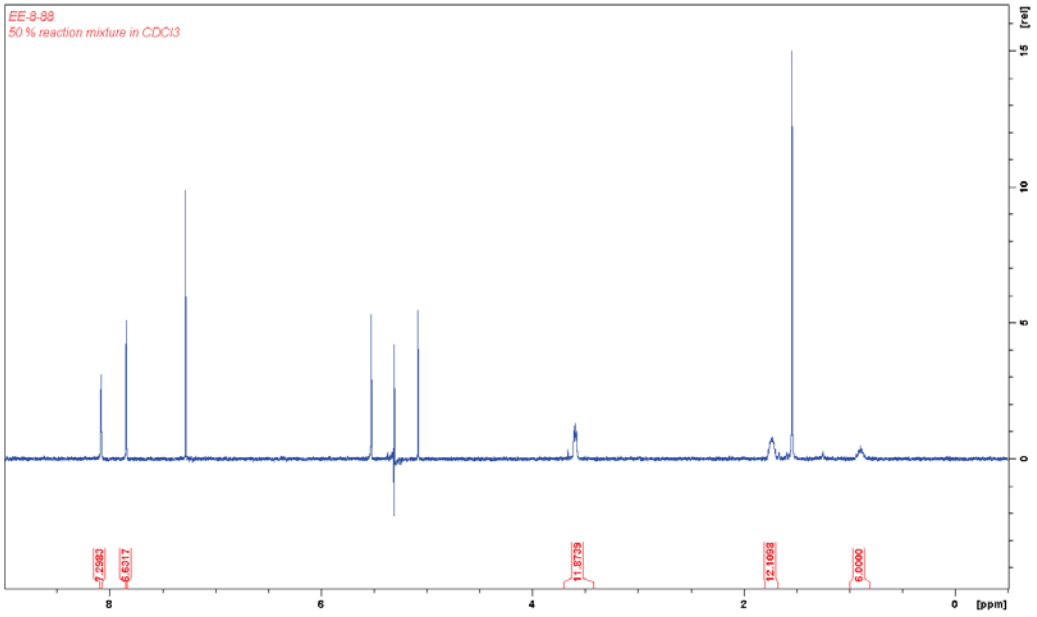


Figure 2.53. ^{1}H NMR spectrum of CC-pentane

Spectra for CC-hexane:

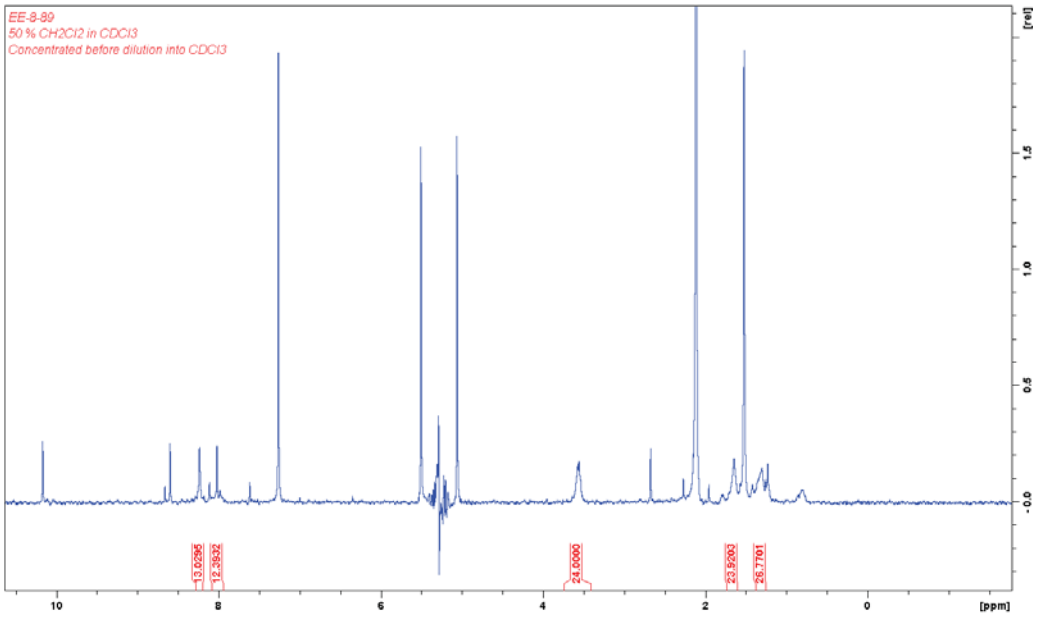


Figure 2.54. ^{1}H NMR spectrum of CC-hexane

Spectra for CC-heptane:

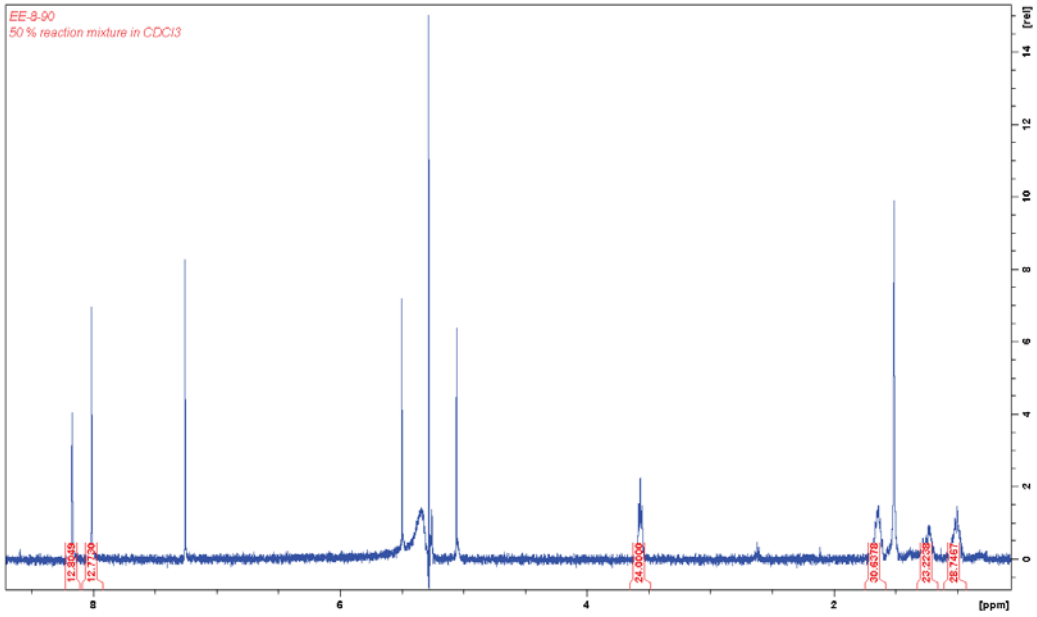


Figure 2.55. ^{1}H NMR spectrum of CC-heptane

Spectra for CC-octane:

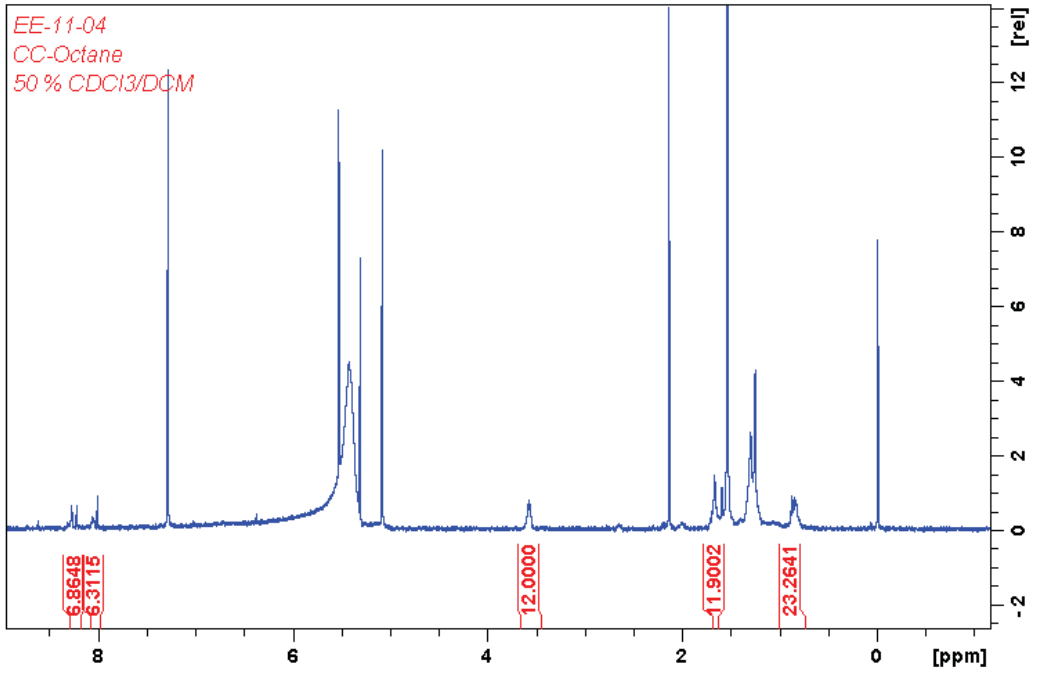


Figure 2.56. ^{1}H NMR spectrum of CC-octane

Spectra for CC-nonane:

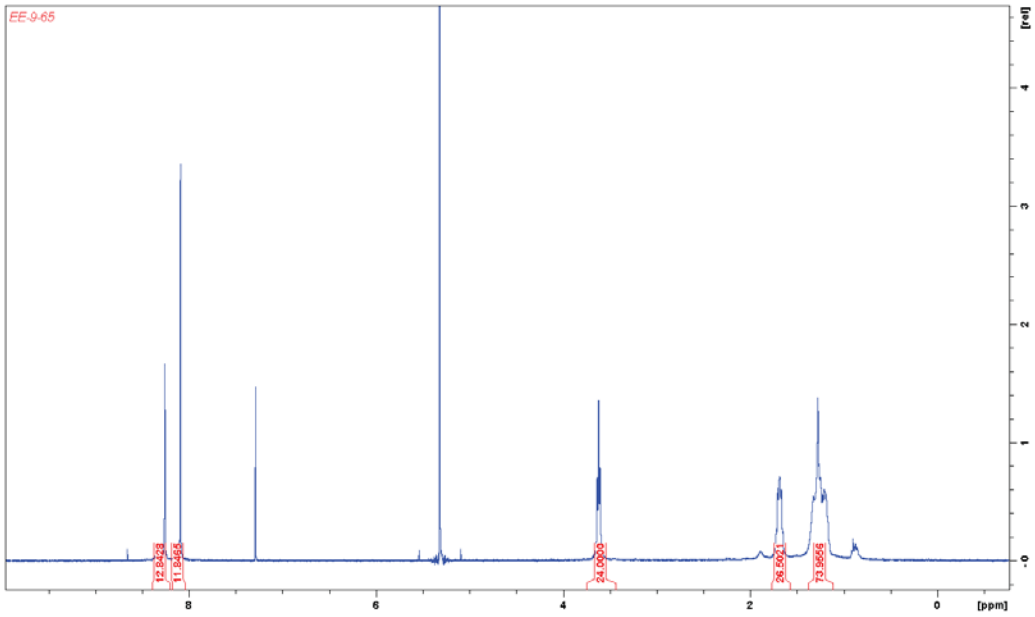


Figure 2.57. ^{1}H NMR spectrum of CC-nonane

2.5 Bibliography

1 S. S. Hidalgo-Tobon, *Concepts Magn. Reson. Part A*, 2010, **36A**, 223–242.

2 M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, John Wiley & Sons, Chichester, 2nd Editio., 2008.

3 G. A. Webb and I. Barsukov, *Nuclear Magnetic Resonance*, Royal Society of Chemistry, Cambridge, 1999, vol. 28.

4 S. V. K. and S. K. Latypov, *Russ. Chem. Rev.*, 2010, **79**, 635.

5 E. Fukushima, *Experimental pulse NMR : a nuts and bolts approach*, Westview Press, Boulder, Co, 1981.

6 K. Song, Q. Bao, F. Chen, C. Huang, J. Feng and C. Liu, *J. Magn. Reson.*, 2016, **268**, 1–9.

7 G. H. Sørland, *Dynamic Pulsed-Field-Gradient NMR*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, vol. 110.

8 C. A. Schalley, Ed., *Analytical Methods in Supramolecular Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2012.

9 C. Schalley, Ed., *Analytical Methods in Supramolecular Chemistry*, Wiley-VCH Verlag GmbH

& Co. KGaA, Weinheim, Germany, 2006.

10 B. D. N. Rao and M. D. Kemple, Eds., *NMR as a Structural Tool for Macromolecules*, Springer US, Boston, MA, 1996.

11 C. F. Tormena, R. Evans, S. Haiber, M. Nilsson and G. A. Morris, *Magn. Reson. Chem.*, 2010, **48**, 550–553.

12 W. Sutherland, *Philos. Mag. Ser. 6*, 1905, **9**, 781–785.

13 A. Einstein, *Ann. Phys.*, 1905, **322**, 549–560.

14 G. T. Evans and C. P. James, *J. Chem. Phys.*, 1983, **79**, 5553.

15 T. Ohtsuki, *J. Chem. Phys.*, 1982, **77**, 1443.

16 M. Stjerndahl, D. Lundberg, H. Zhang and F. M. Menger, *J. Phys. Chem. B*, 2007, **111**, 2008–2014.

17 A. Ortega and J. García de la Torre, *J. Chem. Phys.*, 2003, **119**, 9914.

18 L. Avram and Y. Cohen, *Chem. Soc. Rev.*, 2015, **44**, 586–602.

19 Z. Li and H. Wang, *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.*, 2003, **68**, 61206.

20 Z. J. Derlacki, A. J. Easteal, A. V. J. Edge, L. A. Woolf and Z. Roksandic, *J. Phys. Chem.*, 1985, **89**, 5318–5322.

21 T. Privalov and A. Shalagin, *Phys. Rev. A*, 1999, **59**, 4331–4339.

22 V. L. Saveliev and K. Nanbu, *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.*, 2002, **65**, 51205.

23 B. J. Ackerson, *J. Chem. Phys.*, 1982, **76**, 2675.

24 S. Hanna, W. Hess and R. Klein, *Phys. A Stat. Mech. its Appl.*, 1982, **111**, 181–199.

25 I. Swan, M. Reid, P. W. A. Howe, M. A. Connell, M. Nilsson, M. A. Moore and G. A. Morris, *J. Magn. Reson.*, 2015, **252**, 120–129.

26 A. Jerschow and N. Müller, *J. Magn. Reson.*, 1997, **125**, 372–375.

27 R. SIMHA, *Trans. N. Y. Acad. Sci.*, 1949, **11**, 96–96.

28 G. Jones and S. K. Talley, *J. Am. Chem. Soc.*, 1933, **55**, 624–642.

29 A. J. Masters and T. Keyes, *J. Stat. Phys.*, 1985, **39**, 215–239.

30 M. Li, G. W. Mulholland and M. R. Zachariah, *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.*, 2014, **89**, 22112.

31 E. D. Becker, *High Resolution NMR*, Elsevier, 2000.

32 M. P. Foster, C. A. McElroy and C. D. Amero, *Biochemistry*, 2007, **46**, 331–340.

33 G. C. Levy, A. Kumar and D. Wang, *J. Am. Chem. Soc.*, 1983, **105**, 7536–7540.

34 V. P. Arkhipov, Z. S. Idiyatullin, E. F. Potapova, O. N. Antzutkin and A. V Filippov, *J. Phys. Chem. B*, 2014, **118**, 5480–5487.

35 N. Shemesh, E. Özarslan, A. Bar-Shir, P. J. Basser and Y. Cohen, *J. Magn. Reson.*, 2009, **200**, 214–225.

36 B. Lonetti, A. Tsigkri, P. R. Lang, J. Stellbrink, L. Willner, J. Kohlbrecher and M. P. Lettinga, *Macromolecules*, 2011, **44**, 3583–3593.

37 C. Wang, H. Ding, Y. Yang, B. Li, F. Pan, G. Zhu, M. Zeller, D. Yuan and C. Wang, *Chem. Commun.*, 2014, **51**, 1976–1979.

38 M. Tominaga, K. Suzuki, M. Kawano, T. Kusukawa, T. Ozeki, S. Shakamoto, K. Yamaguchi and M. Fujita, *Angew. Chem. Int. Ed.*, 2004, **43**, 5621–5625.

39 Q.-F. Sun, J. Iwasa, D. Ogawa, Y. Ishido, S. Sato, T. Ozeki, Y. Sei, K. Yamaguchi and M. Fujita, *Science*, 2010, **328**, 1144–1147.

40 T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S. Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner and A. I. Cooper, *Nat. Mater.*, 2009, **8**, 973–978.

41 S. I. Swamy, J. Bacsa, J. T. a Jones, K. C. Stylianou, A. Steiner, L. K. Ritchie, T. Hasell, J. a Gould, A. Laybourn, Y. Z. Khimyak, D. J. Adams, M. J. Rosseinsky and A. I. Cooper, *J. Am. Chem. Soc.*, 2010, **132**, 12773–12775.

42 M. Liu, M. A. Little, K. E. Jelfs, J. T. A. Jones, M. Schmidtmann, S. Y. Chong, T. Hasell and A. I. Cooper, *J. Am. Chem. Soc.*, 2014, **136**, 7583–7586.

43 K. E. Jelfs, E. G. B. Eden, J. L. Culshaw, S. Shakespeare, E. O. Pyzer-Knapp, H. P. G. Thompson, J. Bacsa, G. M. Day, D. J. Adams, A. I. Cooper, O. Edward, H. P. G. Thompson, J. Bacsa, G. M. Day, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2013, **135**, 9307–9310.

44 K. E. Jelfs, X. Wu, M. Schmidtmann, J. T. A. Jones, J. E. Warren, D. J. Adams and A. I. Cooper, *Angew. Chem. Int. Ed.*, 2011, **50**, 10653–10656.

45 D. Holden, K. E. Jelfs, A. I. Cooper, A. Trewin and D. J. Willock, *J. Phys. Chem. C*, 2012, **116**, 16639–16651.

46 J. L. Culshaw, G. Cheng, M. Schmidtmann, T. Hasell, M. Liu, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2013, **135**, 10007–10010.

47 B. N. Solomonov, I. A. Sedov and A. A. Akhmadiyarov, *J. Phys. Org. Chem.*, 2009, **22**, 1142–1147.

48 F. Eblinger and H.-J. Schneider, *J. Phys. Chem.*, 1996, **100**, 5533–5537.

49 V. Santolini, G. A. Tribello and K. E. Jelfs, *Chem. Commun.*, 2015, **51**, 15542–15545.

50 N. R. Voss and M. Gerstein, *Nucleic Acids Res.*, 2010, **38**, W555–W562.

51 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.*, 2005, **167**, 103–128.

52 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

53 J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2007, **127**, 114105.

54 S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B*, 1996, **54**, 1703–1710.

55 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

56 J. N. Israelachvili, in *Intermolecular and Surface Forces*, Elsevier, 2011, pp. 107–132.

57 M. E. Briggs, K. E. Jelfs, S. Y. Chong, C. Lester, M. Schmidtmann, D. J. Adams and A. I. Cooper, *Cryst. Growth Des.*, 2013, **13**, 4993–5000.

58 M. L. Connolly, *J. Appl. Crystallogr.*, 1983, **16**, 548–558.

59 A. Jerschow and N. Müller, *J. Magn. Reson.*, 1997, **125**, 372–375.

60 A. Jerschow and N. Müller, *J. Magn. Reson. Ser. A*, 1996, **123**, 222–225.

61 J. E. Baschek, H. C. R Klein and U. S. Schwarz, *BMC Biophys.*, 2012, **5**, 22.

62 G. Rapenne, C. Dietrich-Buchecker and J.-P. Sauvage, *J. Am. Chem. Soc.*, 1999, **121**, 994–1001.

63 J.-F. Ayme, J. E. Beves, D. A. Leigh, R. T. McBurney, K. Rissanen and D. Schultz, *Nat. Chem.*, 2012, **4**, 15–20.

64 M. Feigel, R. Ladberg, S. Engels, R. Herbst-Irmer and R. Fröhlich, *Angew. Chem. Int. Ed.*, 2006, **45**, 5698–5702.

65 O. Safarowsky, M. Nieger, R. Fröhlich and F. Vögtle, *Angew. Chem. Int. Ed.*, 2000, **39**, 1616–1618.

66 A. Gafni and Y. Cohen, *J. Org. Chem.*, 1997, **62**, 120–125.

67 R. Wimmer, F. L. Aachmann, K. L. Larsen and S. B. Petersen, *Carbohydr. Res.*, 2002, **337**, 841–849.

68 R. Rymden, J. Carlfors and P. Stilbs, *J. Incl. Phenom.*, 1983, **1**, 159–167.

69 B. Levay, *J. Phys. Chem.*, 1973, **77**, 2118–2121.

70 D. J. Cram and J. M. Cram, *Container Molecules and Their Guests*, Royal Society of Chemistry, Cambridge, 1997.

71 N. O'Reilly, N. Giri and S. L. James, *Chem. - A Eur. J.*, 2007, **13**, 3020–3025.

72 N. Giri, M. G. Del Pópolo, G. Melaugh, R. L. Greenaway, K. Rätzke, T. Koschine, L. Pison, M. F. C. Gomes, A. I. Cooper and S. L. James, *Nature*, 2015, **527**, 216–220.

73 J. R. Holst and A. I. Cooper, *Adv. Mater.*, 2010, **22**, 5212–5216.

74 A. I. Cooper, *Angew. Chem. Int. Ed.*, 2012, **51**, 7892–7894.

75 M. Mastalerz, *Chem. - A Eur. J.*, 2012, **18**, 10082–10091.

76 P. Mateus, R. Delgado, P. Brandão and V. Félix, *J. Org. Chem.*, 2009, **74**, 8638–46.

77 S. Kubik, *Angew. Chem. Int. Ed.*, 2009, **48**, 1722–5.

78 P. Mateus, R. Delgado, P. Brandão, S. Carvalho and V. Félix, *Org. Biomol. Chem.*, 2009, **7**, 4661–73.

79 Y. Jin, B. A. Voss, A. Jin, H. Long, R. D. Noble and W. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 6650–6658.

80 J. T. A. Jones, T. Hasell, X. Wu, J. Bacsa, K. E. Jelfs, M. Schmidtmann, S. Y. Chong, D. J. Adams, A. Trewin, F. Schiffman, F. Cora, B. Slater, A. Steiner, G. M. Day and A. I. Cooper, *Nature*, 2011, **474**, 367–371.

81 M. Mastalerz, *Angew. Chem. Int. Ed.*, 2010, **49**, 5042–5053.

82 M. Mastalerz, *Synlett*, 2013, **24**, 781–786.

83 G. R. Desiraju, *J. Mol. Struct.*, 2003, **656**, 5–15.

84 F. C. Pigge, *CrystEngComm*, 2011, **13**, 1733.

85 S. Jiang, J. T. A. Jones, T. Hasell, C. E. Blythe, D. J. Adams, A. Trewin and A. I. Cooper, *Nat. Commun.*, 2011, **2**, 207.

86 R. L. Greenaway, D. Holden, E. G. B. Eden, A. Stephenson, C. W. Yong, M. J. Bennison, T. Hasell, M. E. Briggs, S. L. James and A. I. Cooper, *Chem. Sci.*, 2017.

87 M. A. Little, J. Donkin, J. Fisher, M. A. Halcrow, J. Loder and M. J. Hardie, *Angew. Chem. Int. Ed.*, 2012, **51**, 764–6.

88 K. E. Chaffee, H. A. Fogarty, T. Brotin, B. M. Goodson and J.-P. Dutasta, *J. Phys. Chem. A*, 2009, **113**, 13675–84.

89 T. M. Hermans, M. A. C. Broeren, N. Gomopoulos, P. van der Schoot, M. H. P. van Genderen, N. A. J. M. Sommerdijk, G. Fytas and E. W. Meijer, *Nat. Nanotechnol.*, 2009, **4**, 721–726.

90 D. R. Burfield, R. H. Smithers and A. S. C. Tan, *J. Org. Chem.*, 1981, **46**, 629–631.

91 A. S. Altieri, D. P. Hinton and R. A. Byrd, *J. Am. Chem. Soc.*, 1995, **117**, 7566–7567.

92 D. H. Wu, A. D. Chen and C. S. Johnson, *J. Magn. Reson. Ser. A*, 1995, **115**, 260–264.

93 W. C. Still, M. Kahn and A. Mitra, *J. Org. Chem.*, 1978, **43**, 2923–2925.

Chapter 3. Functionalisation of organic cages

3.1 Introduction

Porous organic cages (see Chapter 1, Section 1.2) are frequently soluble,[1] which can facilitate casting and solution-processing. [2,3] This is important in the production of membranes[2] and coatings.[3] Conversely, MOFs require significant post-synthetic adaptation to become processable,[4–7] and other classes of porous materials, such as zeolites are insoluble under conditions that are consistent with industrial application.[8] Several porous polymers have been developed, which retain processability after polymerisation,[9,10] and derive their porosity from the inflexibility of the polymer backbone.[10] Fine structural control and pore-placement are frequently important in tuning materials for specific function,[11] such as where cooperative host-guest interactions are important to function.[2,12] Thus, the integration of porous organic molecules into functional three-dimensional structures remains a significant goal, as it facilitates the design of cage-cage spacing, but retains solution-processability.

Although integration of cages into MOFs[13] and polymer membranes[14] had been attempted, each presented significant problems. In the first case, **RCC1** (see Chapter 1, Section 1.2) was incorporated into a MOF. This enabled cage-cage spacing to be designed, but rendered the material insoluble.

In the second case, **CC3** (see Chapter 1, Section 1.2) was incorporated into a polymeric membrane. While this material was highly processable, cage crystals were intercalated into the polymer with an average crystal size of 90 nm, with no control over distribution.[14] As such, inter-cage spacing was not programmable, but depended directly on crystal structure.

Creating a desymmetrised cage that could be programmably incorporated into a three-dimensional structure would generate predictable pore locations in a processable

material. We hypothesised that this could be accomplished by creating cages with a polymerisable 'handle' that could be used to link it to a polymeric chain (Figure 3.1).

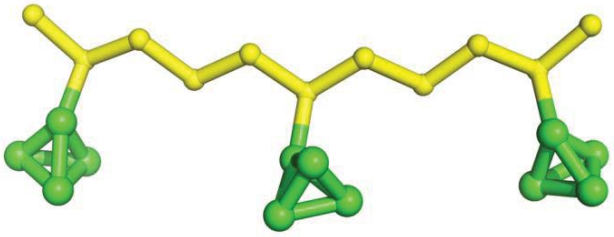


Figure 3.1. Conceptual drawing of cages (green tetrahedra) used as monomers attached to a polymeric backbone (yellow).

In Section 3.2.2, an attempt is made to kinetically-limit the reactivity of cage amine groups in order to produce materials that are functionalised at only one position. In Section 3.2.5, this process is refined, using materials that have already been desymmetrised to produce a single reactive site.

3.1.1 Problems arising from high cage symmetry

Stemming from the symmetry of their synthetic precursors, organic molecular cages are usually highly symmetrical, frequently adopting D_{3h} and T_h symmetries (Figure 3.2).

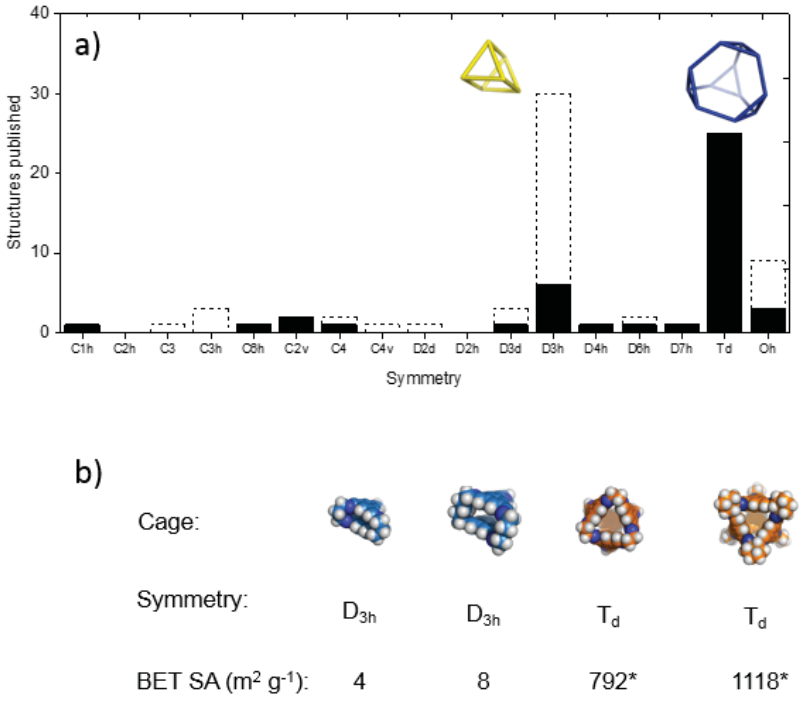


Figure 3.2 (a) Porous organic cages synthesised to date, ordered by symmetry. Filled bars indicate cages where the SA was reported. Dotted bars represent cages where no SA was reported; (b) Four cages, CC-propane and CC-pentane (blue), reported in Section 2.2.1.4, and CC1 and CC3 (orange), demonstrate markedly different porosities.

Many organic cages are synthesised from aldehydes and amines, and have a number of imine groups in the structure. ^{1}H NMR spectra of imine cages usually demonstrate that imine environments are equivalent, which is observed as a single peak associated with the *H*C=N proton (see Chapter 2, Figure 2.39)[15,16] Thus, functionalisation of these cages is generally performed globally, for example by reducing all imines concurrently.[13]

Functional moieties have previously been incorporated into cage structures to allow structural diversification in the presence of imine bonds;[17] these groups are symmetry-related within the structure, and where they have been functionalised, such reactions have been completed globally (Figure 3.3).[13,17,18]

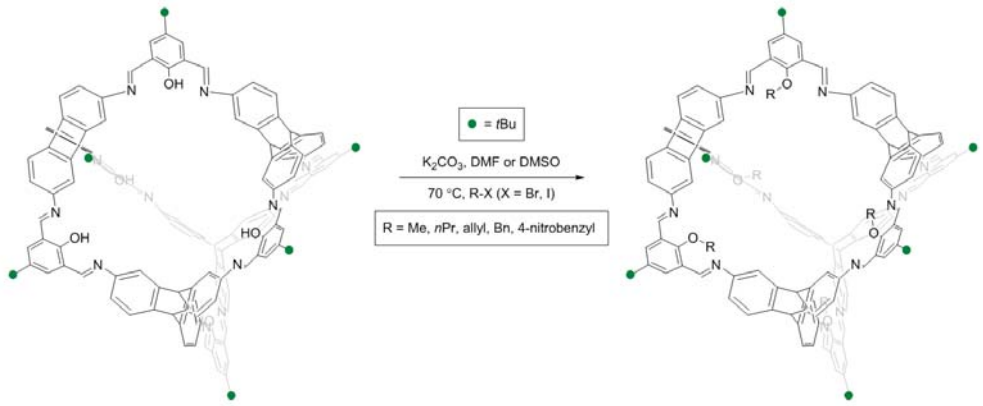


Figure 3.3. Reaction scheme for chemical diversification of organic cages.[17]

In Section 3.2, several attempts are made to desymmetrise organic covalent cages through selective reactions of amine groups.

3.1.2 Attempts to globally functionalise cages

Covalent organic cages were previously functionalised by Culshaw et al. through reduction of **CC1** to form **RCC1**[13] and subsequent reaction with a variety of acyl chlorides, forming a series of dodecaamide cages, **RC1a–d** (Figure 3.4).[18]

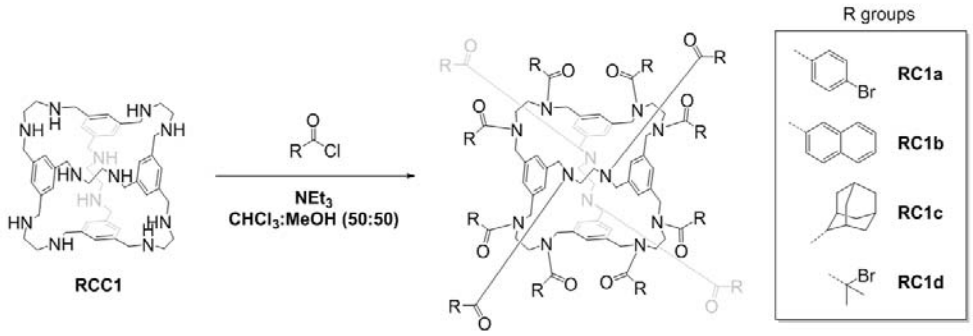


Figure 3.4. Reaction scheme for the functionalisation of RCC1.

The flexibility of **CC1** had previously been found to facilitate reactions in which all 12 nitrogen atoms react externally, forcing the ethylene carbon chain to inhabit the centre of the cage structure in the Cage MOF synthesised by Swamy et al. (Figure 3.5a).[13] Likewise, **RC1a–d** contain inverted ethylenediamine vertices, which may reduce the size of the cage cavity. To determine whether dodecaamide cages would retain an internal cage pore, the crystal structure of **RC1a** (Figure 3.5)[18] was analysed to determine the size of the cavity within the cage. This was determined by placing a sphere in the centre of the cage and inflating it until it came into contact with an atom. The size of the cavity was found to be 3.3 Å. Small pores have previously found to be useful in kinetic quantum sieving.[19]

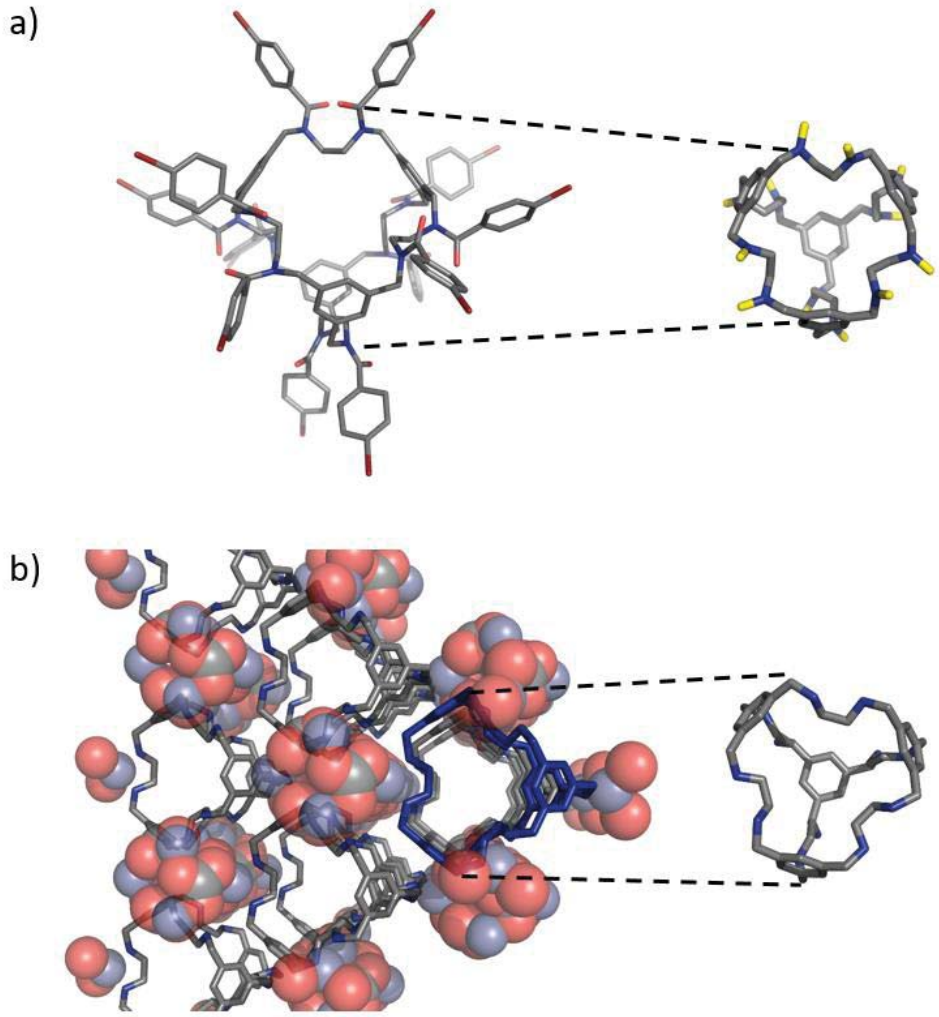


Figure 3.5. (a) The structure of RC1a using single crystal data from reference [18]. Amides are drawn as yellow sticks (right) for clarity; (b) The structure of the cage MOF in reference [13] containing RCC1 and zinc clusters. One RCC1 molecule has been highlighted blue. This molecule has been extracted by deleting all other atoms and is shown on the right. Hydrogens are omitted for clarity.

RC1e, the iodo-analogue of **RC1a**, was subsequently extended *via* palladium coupling to form dendrimers **RC1f–h**. **RC1a–h** synthesised by addition of an excess of 1.2 eq. of acyl chloride per amine in the presence of the non-nucleophilic base NEt_3. Full substitution was attributed to the low steric demands and high structural flexibility of **RCC1**.[18]

3.1.3 Attempts to capitalise on low symmetry

Cage scrambling,[20] which is explored in Section 3.2.5, can significantly decrease the symmetry of cages, by incorporating more than one diamine molecule into the cage framework. Scrambled cages, which consist of a mixture of desymmetrised isomers, have been reported as a method by which to increase surface area.[20] This was accomplished by mixing the starting materials of two cages together in a single pot (Figure 3.6). Benzene-1,3,5-tricarboxaldehyde, which had been found to show no preference in dynamic reactions with more than one vicinal diamine, reacts stochastically with *R,R*-cyclohexane diamine and ethylenediamine. This generates a family of cages (ten possible isomers) with mixed make-up.

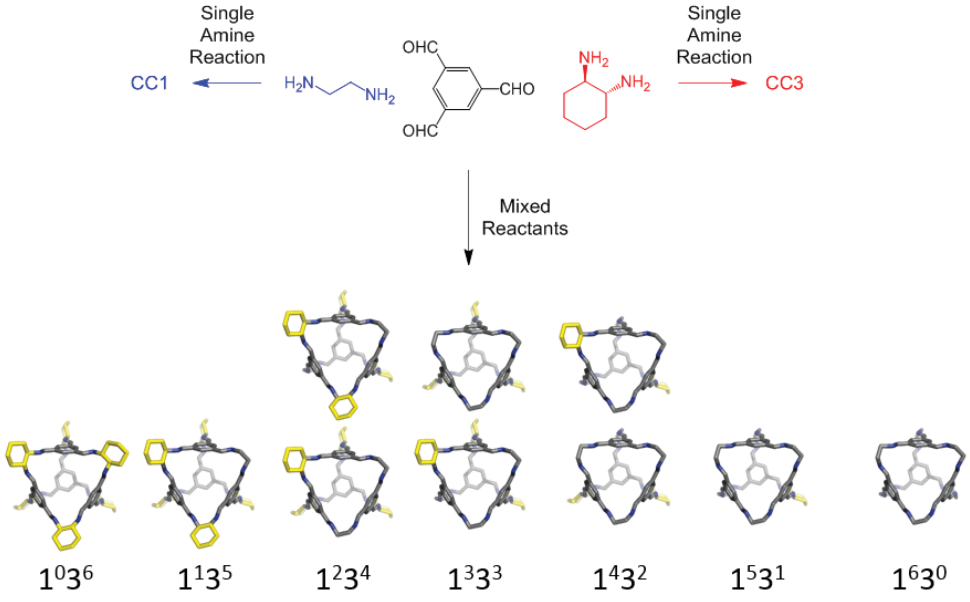


Figure 3.6. A process of reacting mixed amine species with a single trialdehyde was used to generate a mixture of desymmetrised cages. Single-amine reactions are shown in blue (CC1) and red (CC3). In the mixed reaction, cyclohexanediamine vertices are highlighted in yellow.

The solubility of mixtures generated from a 1:1 mixture of amines was found to be highest. This correlated well with the prediction that a 1:1 mixture would generate the

largest number of isomers. Additional positional isomerism was found to increase solubility in chloroform to 337 mg cm^{-3}, compared to CC3, which has a chloroform solubility of 150 mg cm^{-3}. This is attributed to decreased solid-state interactions between positionally desymmetrised cages.

The symmetry of molecular cages, which are traditionally synthesised in one of two topologies (tetrahedral and prismatic, see Figure 3.2) can act to reduce their solubility, which impacts the uptake in a porous liquid.

Prismatic cages are often published in the absence of sorption data (Figure 3.2a). Reasons for this are not clear, but may also stem from their use as receptors, rather than sorbents.[21–24] In addition to this, flexibility of the prismatic structure can also lead to collapse, where a structure with no permanent cavity is more energetically favourable.[25] This decreases sorption effectiveness and may provide a disincentive for investigating sorbents with this molecular shape. Frequently, these cages are also smaller, which may also contribute to a lack of accessible pores.

However, coupled with high symmetry, rigidity can result in high intermolecular interactions, high degrees of crystallinity, and limited solubility.[26,27]

3.2 Results and discussion

To extend the work of Culshaw et al.,[18] and probe the additional chemical space offered by amine functionalisation, a series of acyl chlorides was selected, which targeted the ability to incorporate these moieties into larger functional networks. The selection (Figure 3.7) was designed to probe a variety of steric demands, but primarily focussed on facilitating functional materials. Nicotinoyl chlorides, thiophenes and furans were chosen because of their potential ability to extend functionality beyond the cage *via* further reaction.

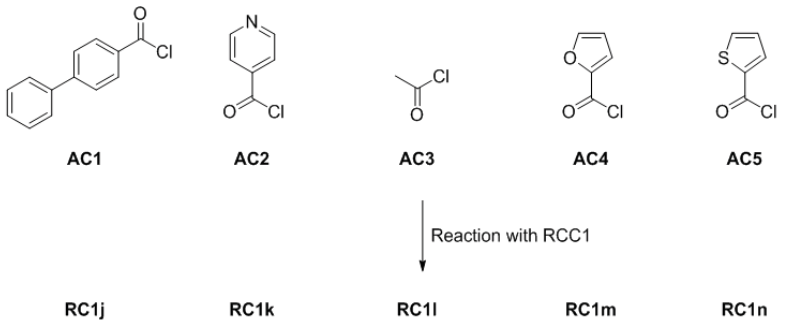


Figure 3.7. Acyl chlorides used in reactions with RCC1 to afford dodecaamide cages

*3.2.1 Synthesis of dodecaamide cages from **RCC1***

Dodecaamide cages **RC1j–n** were synthesised by addition of a slight excess of 1.2 equivalents of acyl chloride to a solution of **RCC1** NEt_3 (1.4 eq. per cage) in dry chloroform, following a procedure modified from that of Culshaw et al.[18] ^{1}H NMR chemical shifts at 7.1, 3.7 and 2.7 ppm, associated with **RCC1** were observed to broaden on reaction, which was consistent with data already published. This was rationalised by Culshaw et al. to be a result of limited rotation about the amide bond, which had resulted in a large number of rotamers present in the structure.[18]

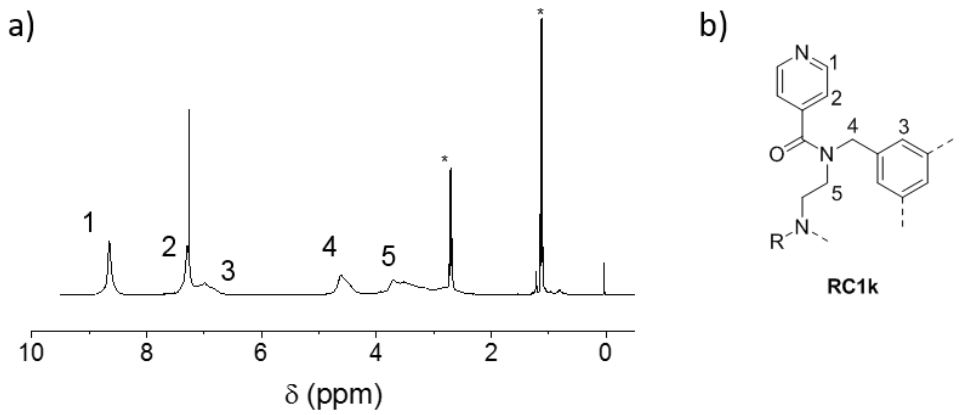


Figure 3.8. (a) ^{1}H NMR spectrum of RCC1j. Asterisks refer to peaks associated with NEt_3. (b) A fragment of the chemical structure of RCC1j. Numbers refer to chemical shift assignment.

3.2.2 Characterisation of dodecaamide cages

An excess of acyl chloride was found to be essential in generating fully substituted cages. Furthermore, cages that have not been fully functionalised, which contain secondary amine moieties with a significantly lower pK_a than the tertiary amides,[28,29] were found to ionise much more strongly. These peaks were found to dominate mass spectra in which cages were not fully substituted, even where ^{1}H NMR analysis indicated by integral area that average number of amines that had reacted approached 12 (Figure 3.9).

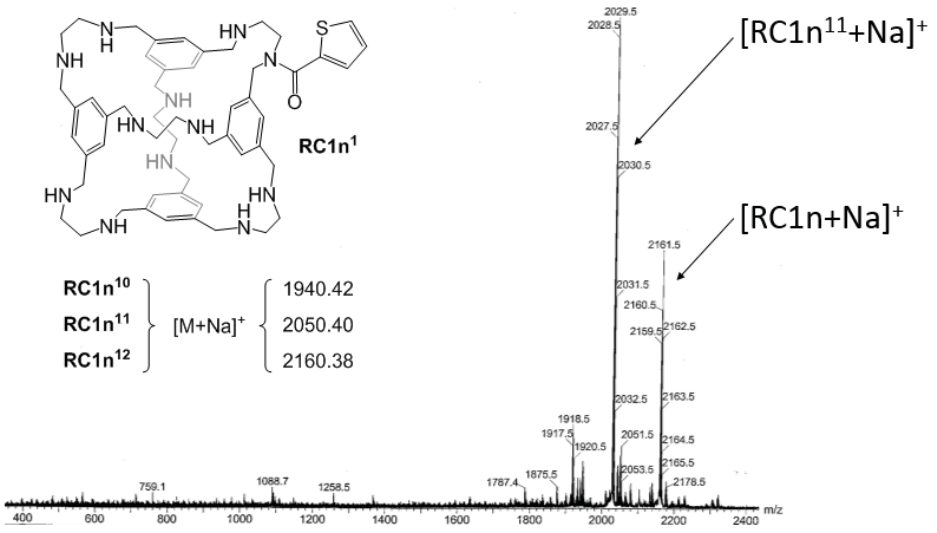


Figure 3.9. Representative mass spectra for reactions between RCC1 and 12.0 eq, of AC5 per cage

Reactions between **RCC1** and an excess of 1.2 equivalents of **AC2–5** per cage in the presence of an excess of 1.4 equivalents of NEt_3 per cage were analysed by high resolution mass spectrometry, and were found to produce spectra with a base peak (the most abundant peak) associated with a 12-substituted product (Figure 3.10).

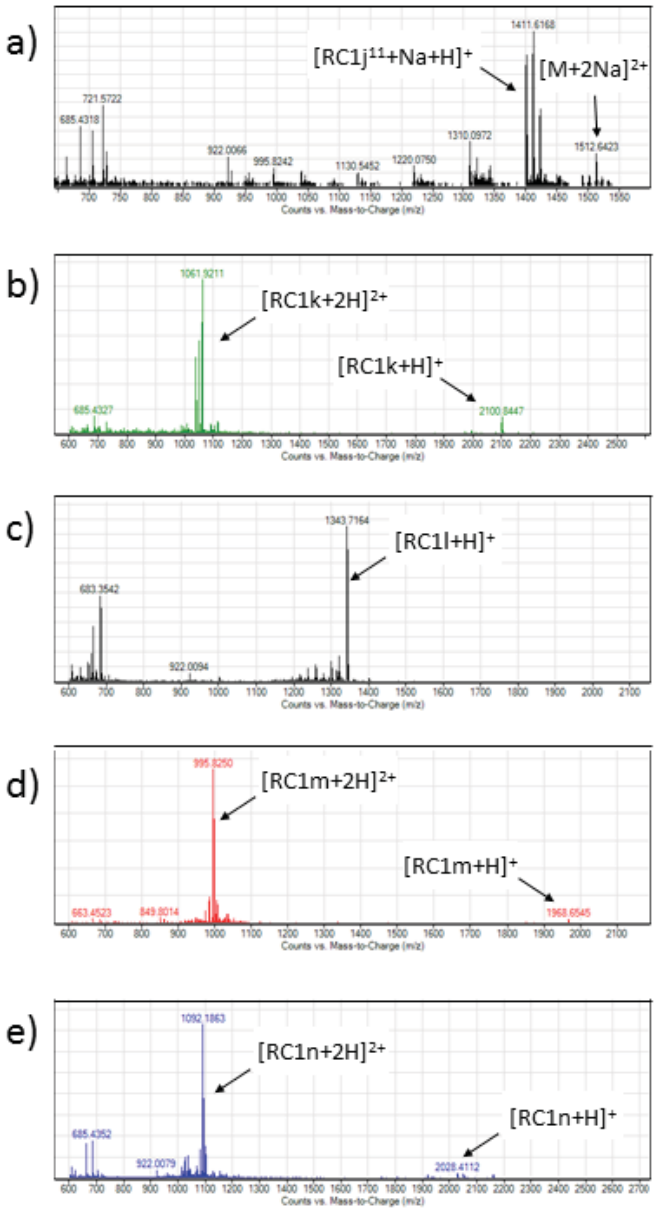


Figure 3.10. High resolution mass spectrum for the products of reactions between RCC1 and AC1–5. (a) RC1j; (b) RC1k; (c) RC1l; (d) RC1m; (e) RC1n.

Reactions between **RCC1** and the most sterically-demanding acyl chloride, **AC1**, generated mass ions indicative of an 11-substituted cage. Mass spectra demonstrated a single peak at 1400.6256 *m/z*, which corresponded to a doubly-charged 11-substituted species [**RCC1j**^{11}+2H]$^{2+}$ (1400.6271 *m/z* calculated for $C_{191}H_{162}N_{12}O_{11}^{2+}$). Analysis by ^{1}H NMR spectroscopy indicated that broadening of chemical shifts at 3.75 and 4.75 ppm was more pronounced than in ^{1}H NMR spectra of **RCC1k–n**. We hypothesised that this may be a result of the lower symmetry associated with an 11-substituted cage. However, this additional broadening made accurate integration difficult, and the ratio between aromatic protons (8.0 to 7.3 ppm) and the broad peak at 4.75 ppm varied significantly according to the integral regions selected (Figure 3.10a).

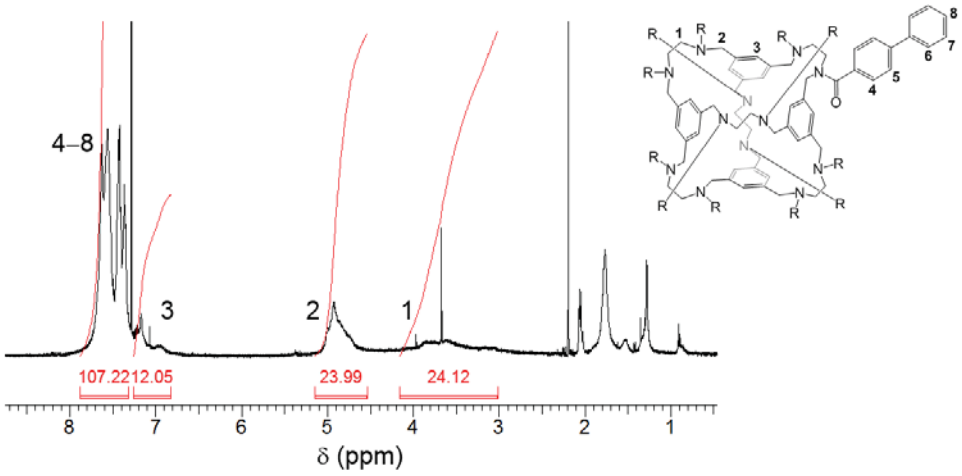


Figure 3.11. ^{1}H NMR spectrum of products associated with the reaction between RCC1 and AC1.

To confirm the structure of **RCC1j–n**, infrared (IR) spectra were recorded and compared to that of **RCC1**. Appearance of a prominent absorption at 1629 cm^{-1}, which is consistent with a C=O stretch,[30] confirmed the formation of amides (Figure 3.12a).

Furthermore, all IR spectra, including that of **RCC1j**, demonstrated a complete loss of absorption at 2813 cm^{-1} and 3260 cm^{-1}, which were present in the IR spectrum of **RCC1** (Figure 3.12). This confirmed that amines had fully reacted to form a dodecaamide cage.

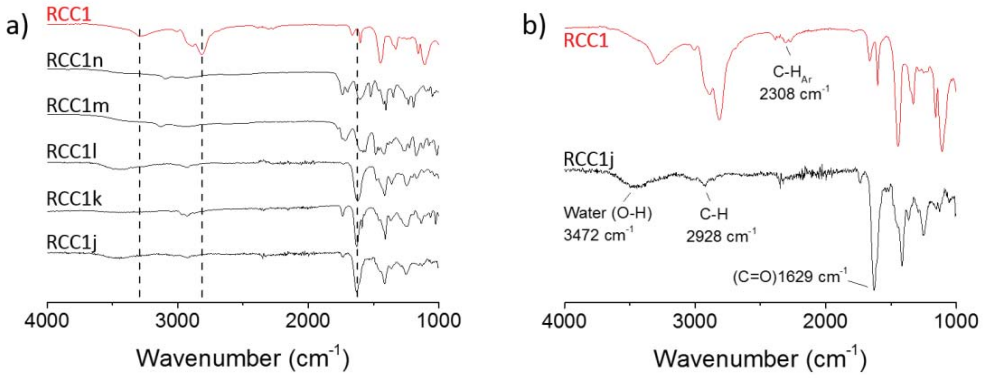


Figure 3.12. (a) Infrared spectra of RCC1 (red) and RCC1j–n (black); (b) Infrared spectra of RCC1 (red) and RCC1j (black) with prominent peaks highlighted.

3.2.3 Attempts to functionalise RCC3 with acyl chlorides

RCC1 had been found to invert ethylenediamine vertices inside the cavity of the cage (Figure 3.5).[13,18] Conversely, we hypothesised that cyclohexane groups in **RCC3** would be unable to be unable to occupy the internal cavity of the cage because their steric bulk was significantly higher. To test this, we evaluated the crystal structure of **RCC3**[15] and comparing the size of the cyclohexane ring with the dimensions of the internal cavity. The cyclohexane ring, which was measured between the external hydrogens, was approximately 4.9 Å wide. The internal cavity was measured as 9.8 Å in radius by placing a sphere in the centre of the cage and expanding it until it came into contact with an atom (Figure 3.13).

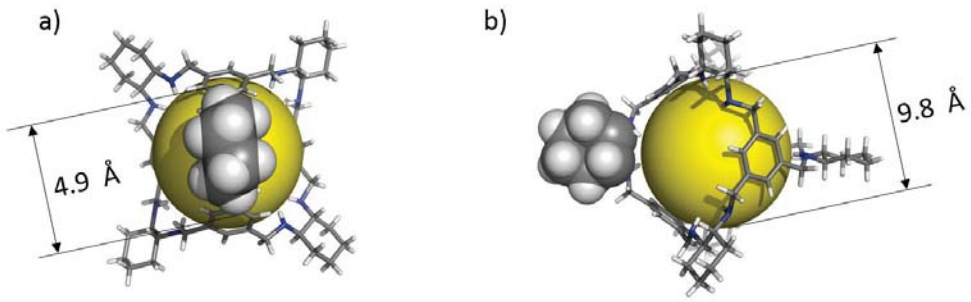


Figure 3.13. Structure of RCC3 viewed (a) along the C_2 axis and (b) from the side. Internal cavity is highlighted with a yellow sphere. One cyclohexane group has been shown as a space-filled structure to help gauge size.

Given the size of the molecular cavity within **RCC3**, we hypothesised that a flexible cage could incorporate a single cyclohexane, and generate a structure with two externally-facing nitrogen atoms. However, it was not clear whether this would be energetically favourable. It was clear from these calculations that no more than one cyclohexane could be incorporated into the cavity, and it was hypothesised that the steric constraints that surrounded amine nitrogens in **RCC3** would affect the degree to which cages could be functionalised. To test this, **RCC3** was exposed to an excess 1.2 equivalents of acyl chlorides **AC1–5** per amine, in the presence of 1.4 equivalents of NEt_3 for every cage. ^{1}H NMR spectroscopic analysis confirmed that, as with reaction between acyl chlorides and **RCC1**, reactions reached completion within minutes. However, average substitutions were observed by mass spectrometric analysis to be significantly below saturation (Figure 3.14).

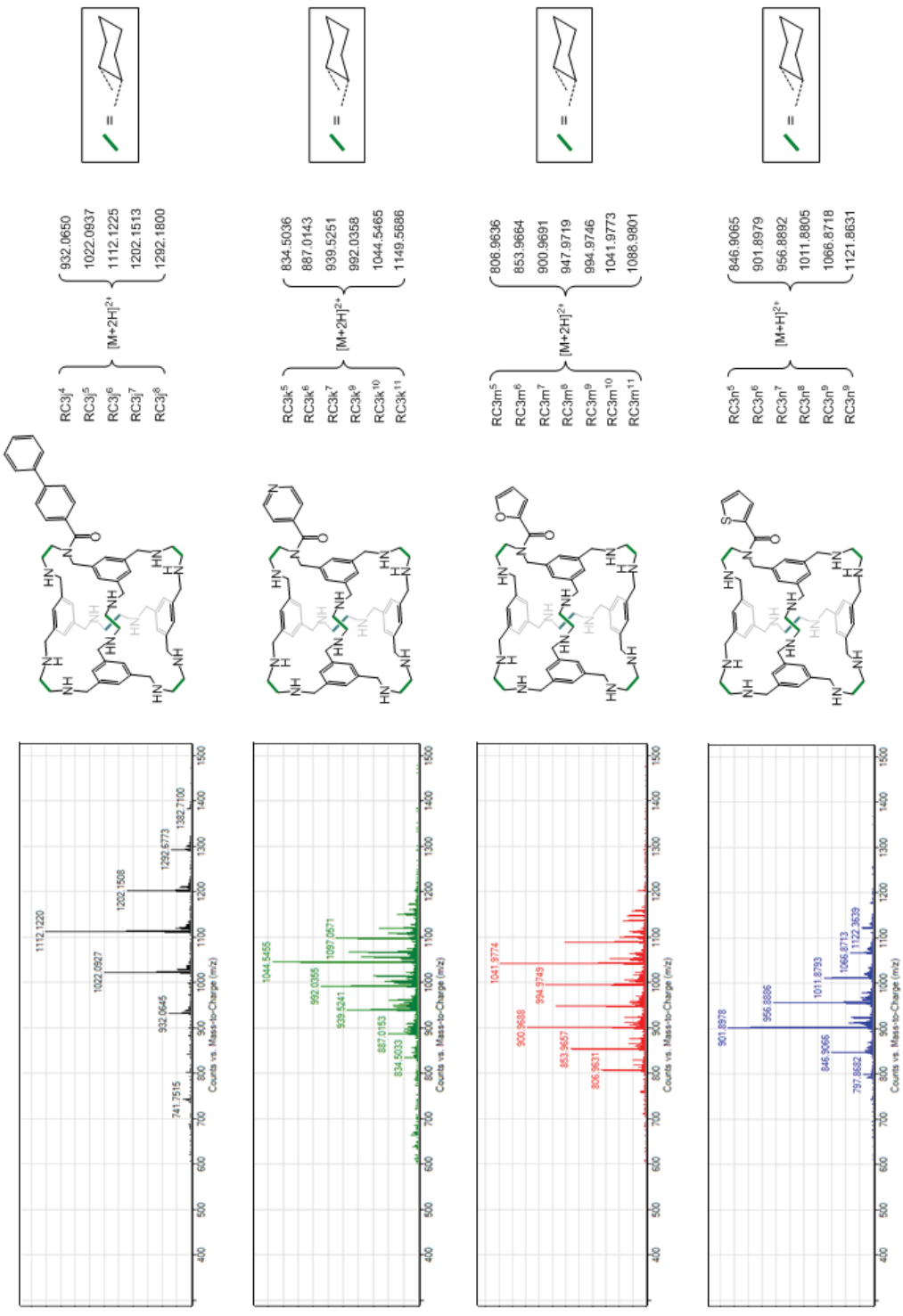


Figure 3.14. (a) High resolution mass spectra of acylated RC3 (left), presented alongside chemical structures of mono-acylated cages and the calculated masses of charged, more highly-substituted species.

Several attempts were made to measure the kinetics of reactions between reduced cages and acyl chlorides. However, at room temperature, reactions between **AC1–5** and both **RCC1** and **RCC3** were found to be complete in less than sixty seconds: this was approximately the minimum time required to lock and shim a sample prior to acquisition within the NMR spectrometer. Subsequent acquisitions were obtained in seconds. In the case of **RCC1**, starting materials were observed to have disappeared before the first acquisition. In the case of **RCC3**, starting materials remained, but no further reaction was noted after the first acquisition.

To investigate the reaction at times less than sixty seconds, a large excess of 120 equivalents of propylamine per cage was added to a reaction between **RCC3** and **AC2** after fifteen seconds. Propylamine is sterically-unhindered, and more nucleophilic than secondary amines.[31] Thus, an excess of propylamine was predicted to quench any remaining acyl chloride, and prevent further reaction with **RCC3**. Crude mixtures were analysed by mass spectrometry (Figure 3.15) and some qualitative evidence of additional reaction was noted between samples that were quenched after 15 seconds and those that had been left for 4 hours. No additional reaction was noted after 4 hours. Products were named **RC3k^{x}**, where x indicates the number of amines substituted.

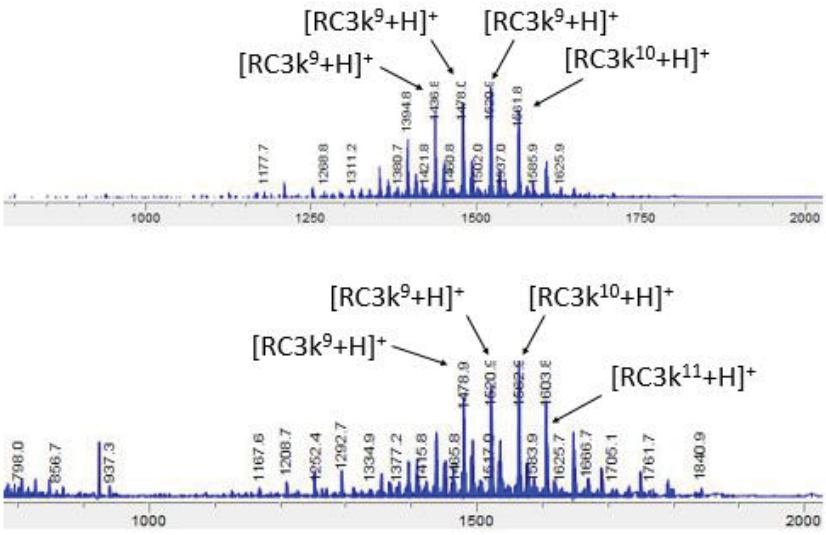


Figure 3.15. Mass spectra of crude mixtures for the reaction between RCC3 and AC3 taken (a) in a reaction that was quenched after fifteen seconds and (b) in a separate reaction that had been allowed to reach completion over 4 hours.

To further investigate reaction kinetics, reactions between **RCC3** and **AC1–5** were cooled to -41 °C in an acetonitrile/CO_2 bath, which was predicted to decrease the rate of reaction. One reaction was quenched after 15 seconds by the addition of 120 equivalents of propylamine per cage, whist a second reaction was removed from the bath and allowed to warm to room temperature over 4 hours. Crude solutions were analysed by mass spectrometry, which confirmed that both unquenched and quenched reactions demonstrated similar reaction profiles (Figure 3.16). At lower temperatures, reactions were observed to generate less-substituted products, even when reactions were allowed to warm to room temperature. Differing substitution levels may be a result of reduced solubility at decreased temperature, although reactions were not observed to contain precipitate. Further work would include investigating differences in solubility by filtering reactions at temperature before allowing crude reaction mixtures to warm to room temperature.

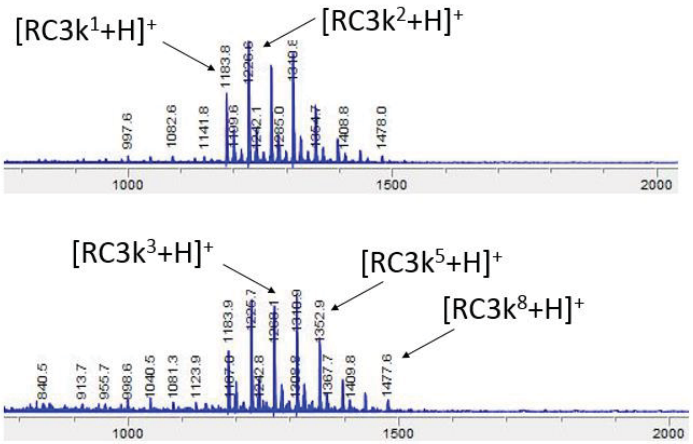


Figure 3.16. Mass spectra of reactions between RCC3 and AC2 that have been (a) matured for 16 hours at room temperature and (b) quenched after 15 seconds at -41 °C.

A second concern in analysing this data was that small differences between reaction profiles could not be quantitatively analysed, as soft ionisation generally affords only qualitative results.[32] Conversely, ^{1}H NMR data can afford quantitative data, but the broadness of chemical shifts caused by rotamers prevented determination of substitution patterns for individual components.

Reaction rates between amines and anhydrides are significantly slower,[33,34] and have previously been measured by ^{1}H NMR spectroscopic analysis over minutes to hours.[35] Thus, in attempting to reduce the rate of reaction between acyl groups and reduced cages, we hypothesised that acyl chlorides converted to acid anhydrides, may afford a more tractable substrate for studying the kinetics of reaction.

3.2.4 Reactions between reduced cages and acid anhydrides

The rate of nucleophilic addition, such as the reaction between **AC1–5** with amines, is frequently limited by the rate of addition to the carbonyl carbon.[36] This is

predominantly mediated by the dipole moment surrounding the carbonyl bond, and the subsequent partial positive charge at the carbonyl carbon atom.[36] In order to decrease the dipole at the carbonyl bond, it is possible to convert acyl chlorides to acid anhydrides through reaction with their parent carboxylic acid (Figure 3.17).

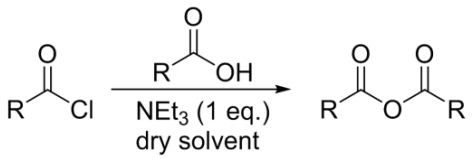


Figure 3.17. General reaction conditions for the synthesis of acid anhydrides from acyl chlorides.

As an initial test to determine whether anhydrides would react more slowly with reduced covalent cages, an excess of 24 equivalents of acetic anhydride was added to a solution of **RCC1** in chloroform-*d*. The resultant reaction was monitored by ^{1}H NMR spectroscopic analysis, with one spectrum acquired every 2 minutes for 30 minutes. As had previously been reported (see Section 3.2.1), ^{1}H NMR chemical shits were observed to broaden even after two minutes of reaction (Figure 3.18).

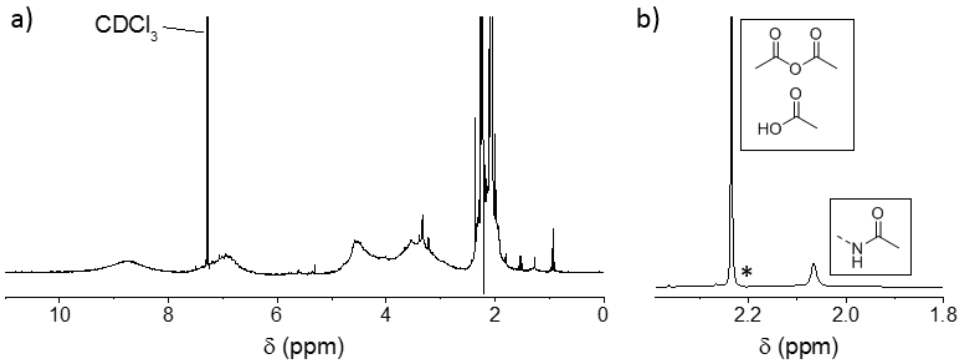


Figure 3.18. (a) ^{1}H NMR spectrum of the reaction between RCC1 and acetic anhydride. Spectrum was acquired after 2 minutes. Peaks between 1.8 and 2.4 ppm have been truncated in height. The vertical line at 2.2 ppm has been caused by a slight distortion, which is highlighted in (b) by an asterisk; (b) The same spectrum focussed around the aliphatic region associated with terminal methyl protons adjacent to the carbonyl carbon.

The peak at 2.06 ppm, which was associated with the terminal methyl protons of acylated cage (Figure 3.18b), was observed to shift upfield over 30 minutes, to 2.09 ppm (Figure 3.19). This was associated with a decrease in pH caused by production of acetic acid during the reaction. Integration of ^{1}H NMR chemical shifts indicated that the integral area between 2.05 and 2.1 ppm only increased by 5 % over the course of measurements. We hypothesised that this was because the reaction between **RCC1** and acetic anhydride was nearing completion before the first acquisition, after approximately 2 minutes. Furthermore, as peaks associated with **RCC1** broadened significantly on acylation, it remained difficult to quantify the degree to which the cage was substituted. To solve this, a calibrated capillary containing tetramethylsilane (TMS) was placed inside the NMR tube in which reactions were studied. This provided a sharp peak at 0 ppm, which did not overlap with ^{1}H NMR chemical shifts associated with the reaction, and would not change over the course of measurements. Details of calibrations can be found in the Experimental, in Section 3.4.

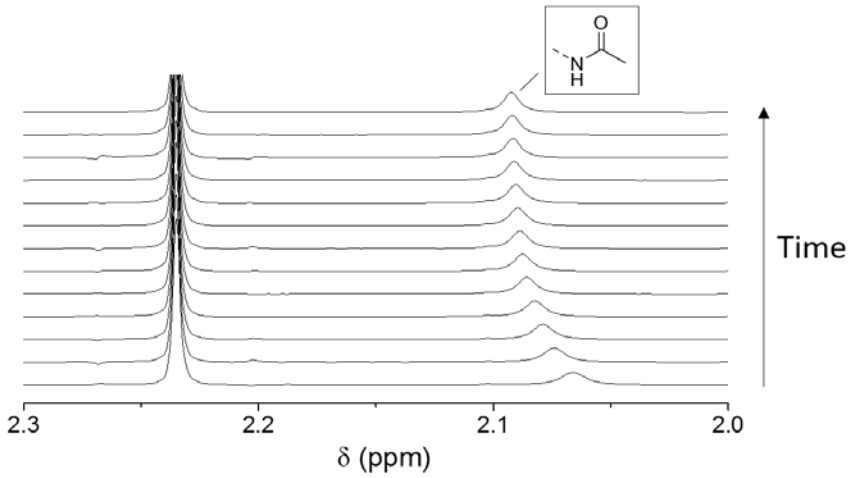
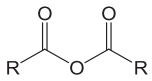


Figure 3.19. ^{1}H NMR spectra of the reaction between RCC1 and acetic anhydride focussed around the aliphatic region associated with terminal methyl protons.

As reactions between **RCC1** and acetic anhydride appeared to be nearing completion within the time taken to lock and shim a ^{1}H NMR sample, it was important to select anhydrides that would enable not only thermodynamic products, but also facilitate measurable real-time analysis of reaction mixtures. To test the effects of steric bulk on reaction rate, a series of commercially available anhydrides was purchased with increasing steric bulk (Table 3.1). We hypothesised that the steric bulk surrounding the carbonyl carbon would also contribute significantly to the rate of reaction, and anhydrides were selected to vary in substitution at the α-carbon. Selecting anhydrides that varied at the α-carbon also ensured an aliphatic chemical shift that was unlikely to overlap with cage peaks.



Entry	R	Abbrev.
1	Me	AA1
2	Et	AA2
3	iPr	AA3
4	tBu	AA4

Table 3.1: Commercially available anhydrides

Using **AA3**, which was predicted to have a significantly lower rate of reaction as a result of its increased steric bulk, several model reactions were conducted to establish the validity of ^{1}H NMR results.

3.2.4.1 Calibrating ^{1}H NMR spectroscopic analysis using a model reaction

Prior to investigations with cages, it was important to test whether conversions could be quantitatively calculated by ^{1}H NMR spectroscopy, and subsequently test whether these measurements could be conducted in real time. To test the degree to reliability of results generated using an internal capillary standard, **AA3** (2.0 μL, 2.5 mmol, 1 eq.) in $CDCl_3$ (0.5 mL) was added to a solution of propylamine (4.2 μL, 2.5 mmol, 1 eq.) in $CDCl_3$ (0.5 mL) and the resultant crude mixture was analysed spectroscopically. Given the high degree of reactivity associated with propylamine, this experiment was predicted to be complete within minutes. Thus, the purpose of this test was to determine whether peak resolution was sufficient to characterise reactions quantitatively, and assess the error associated with these results. Three chemical shifts were identified, which corresponded to remaining excess propylamine (Figure 3.20b,

asterisked peaks). Two further sets of peaks were found: one corresponding to isobutyric acid (Figure 3.20b, peaks 1 and 2); and a second corresponding to the reaction product, N-propylisobutyramide (Figure 3.20b, peaks 4–9).

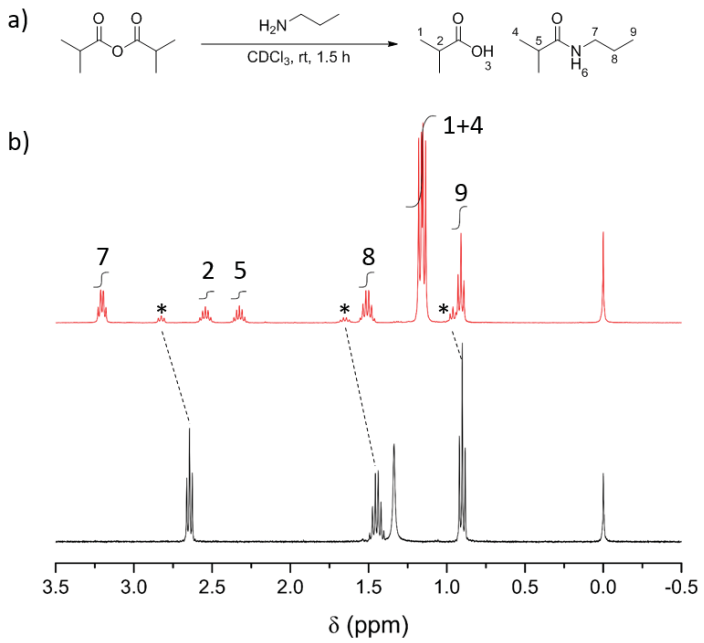


Figure 3.20. (a) Scheme for the reaction between isobutyric anhydride and *n*-propylamine; (b) ^{1}H NMR spectra for *n*-propylamine (black) and the first ^{1}H NMR acquisition after addition of isobutyric anhydride. Asterisked peaks represent residual n-propylamine. Dotted lines are intended to guide the eye between spectra. Numerical labelling in (a) corresponds to chemical shift labelling in (b). ^{1}H NMR chemical shifts of isobutyric acid were assigned by comparison with literature values.[37] ^{1}H NMR chemical shifts for N-propylisobutyramide were assigned by analogy.

In each case, peaks were well resolved. The concentration of N-propylisobutyramide in solution was calculated by measuring the integral area of the chemical shift at 1.5 ppm and comparing it with the integral area of the capillary standard (Figure 3.21a). The chemical shift at 1.5 ppm was selected for comparison as peak overlap with other chemical shifts was minimised. This is important in ensuring

the accuracy of quantitative data, as peak overlap can lead to significant problems in determining the absolute integral size associated with individual components. The concentration of propylamine prior to addition was calculated using a spectrum acquired before the addition of **AA3** and was determined to be 0.0234 mM, and subsequently the concentration of N-propylisobutyramide after the final acquisition was calculated to be 0.0231 mM with a standard deviation of 0.012 mM. This indicated a conversion of 99 ± 1 %. No chemical shifts associated with isobutyric anhydride appeared in ^{1}H NMR spectra of the crude reaction mixture, indicating it had been completely consumed.

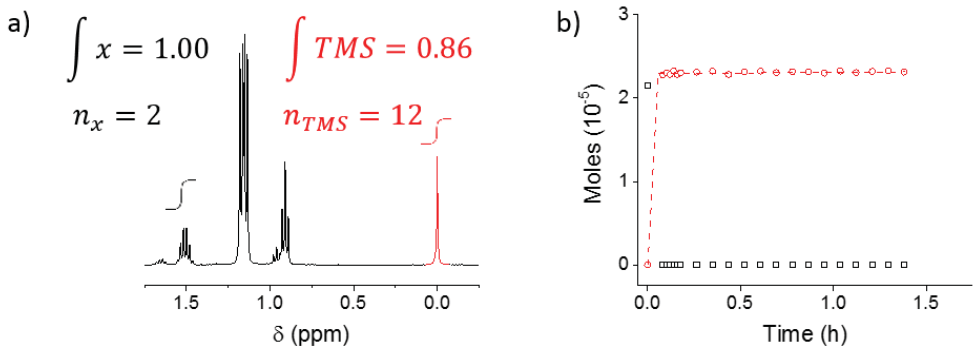


Figure 3.21. (a) The first ^{1}H NMR spectrum, acquired 4 minutes after addition of isobutyric anhydride. Relative integrals of N-propylisobutyramide and TMS are shown. The chemical shift associated with TMS has been highlighted in red; (b) calculated concentrations of N-propylisobutyramide (red) and isobutyric anhydride (black) over time.

With more complex systems, unambiguous determination of chemical shifts can become problematic: where more than one reaction site exists in a molecule, several reaction products are possible. To explore this, isobutyric anhydride was added to a cage-like small molecule analogue, N,N'-dibenzylethane-1,2-diamine (Figure 3.22a). This allowed a controlled comparison to be made with a molecule that structurally

resembled cage, whilst limiting the number of possible reaction products to two (Figure 3.22b–c).

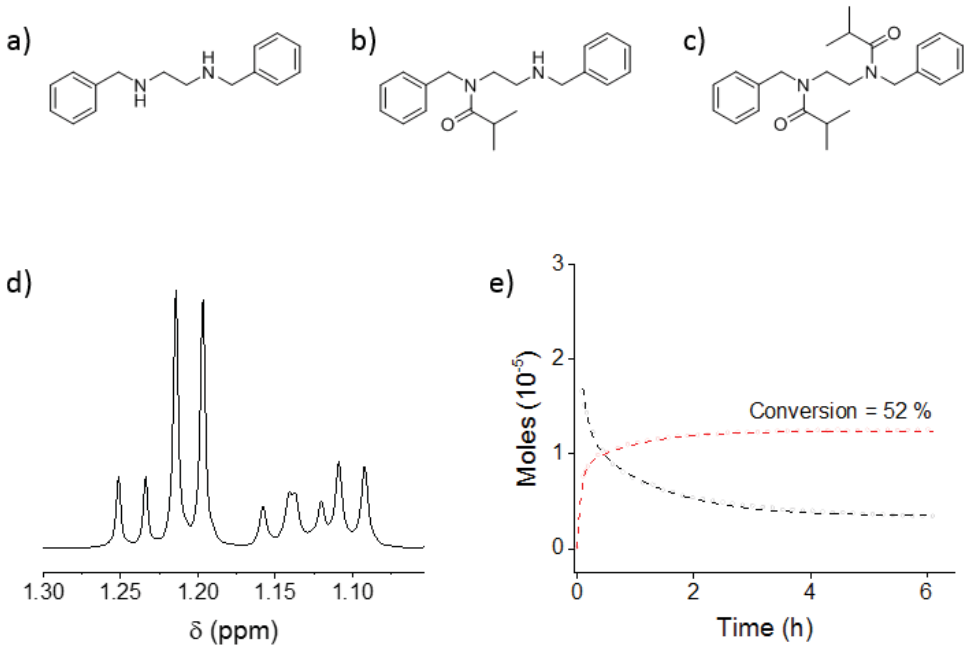


Figure 3.22. (a) N,N'-dibenzylethane-1,2-diamine; (b) The reaction product between N,N'-dibenzylethane-1,2-diamine and a single equivalent of isobutyric anhydride; (c) the double-reaction product between N,N'-dibenzylethane-1,2-diamine and isobutyric anhydride; (d) ^{1}H NMR spectrum of a crude reaction mixture for the reaction of N,N'-dibenzylethane-1,2-diamine and isobutyric anhydride. Highlighted, the chemical shift associated with isobutyric anhydride (blue) and isobutyric acid (red); (e) Moles of isobutyric anhydride (black) and amide products (red) calculated by comparison with TMS internal reference capillary.

As expected, the reaction between isobutyric anhydride and the more sterically-hindered diamine, N,N'-dibenzylethane-1,2-diamine, resulted in a significantly lower rate of reaction, which appeared complete after 5 hours. The multiplet between 1.08 and 1.17 ppm indicated that several products may have been formed (Figure 3.22d). Integration of this entire region was used to calculate conversion, which approached 52 % (Figure 3.22e). This was calculated by comparing the peak integral of the

multiplet, which corresponded to the terminal methyl protons of reaction products, with that of the internal reference capillary. To test whether these products may have been the result of protonation in the acidic environment generated by formation of isobutyric acid, the experiment was repeated in the presence of two equivalents of NEt_3 per N,N'-dibenzylethane-1,2-diamine. The rate of reaction was significantly increased, with the reaction approaching completion within 1 hour. However, measuring conversions was severely hampered by chemical shift overall of the product peaks with those of NEt_3 (Figure 3.23). As peak overlap precluded any further analysis, subsequent reactions were conducted in the absence of NEt_3.

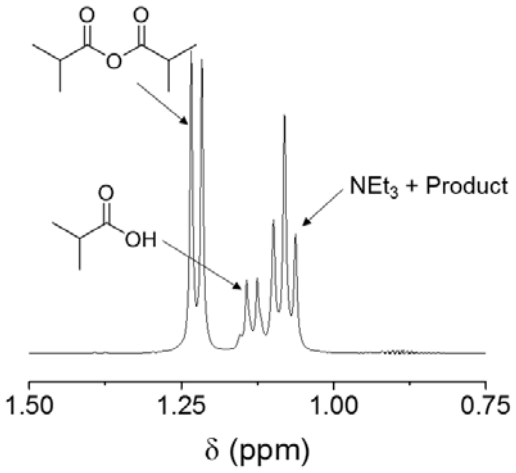


Figure 3.23. First ^1H NMR acquisition for the reaction between N,N'-dibenzylethane-1,2-diamine and AA3 in the presence of NEt_3.

In the absence of NEt_3, the quantity of isobutyric anhydride was found to decrease over 6 hours. From its theoretical starting concentration of 2.52 mM (determined by the quantity of liquid added), it was observed to decrease by 87 % over the course of the reaction (Figure 3.22e). We hypothesised that this excess reactant consumption was the result of side-reactions, such as hydrolysis by adventitious water. To determine whether this apparent mismatch in reactant and product conversions was

the result of reactant consumption, the rate at which consumption of reactants compared with the production of products (Figure 3.24).

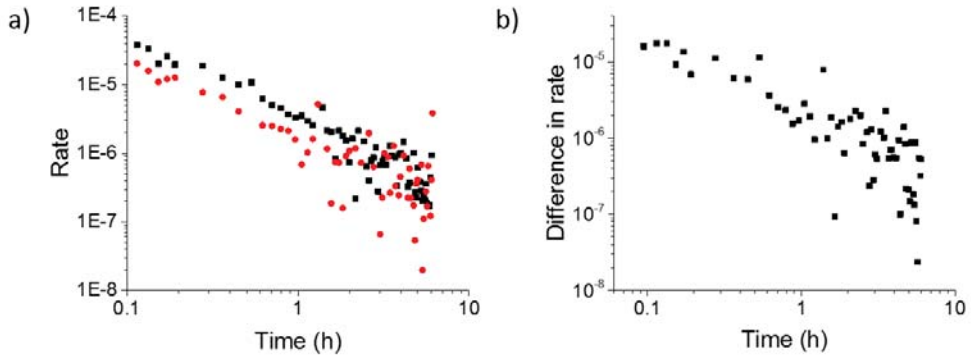


Figure 3.24. (a) Rates of reaction for isobutyric anhydride consumption (black) and product generation (red) are presented on a logarithmic scale; (e) The difference between rates of reaction are displayed on a logarithmic scale.

Reaction rates were calculated by dividing the difference between integral areas of two adjacent points and dividing by the time difference between them. A double-log plot between reaction rates and time facilitated a simple comparison between the rates of reaction, as it produced a straight-line, which was relatively easy to interpret (Figure 3.24a). Here, the magnitude of the rate could be used to estimate the speed with which reactants were consumed and products were formed. Rates of reactant consumption and product generation should be identical where reactants convert only to the products measured. However, the rate of reactant consumption was found to be ~10 % higher than that of product generation (Figure 3.24b). This proportional difference remained constant over time, although both rates were negligible after ~2 hours. This may indicate an additional reaction, such as hydrolysis, although product peaks between 1.08 and 1.18 ppm initially overlapped with the doublet associated with isobutyric acid. This doublet moved downfield over 6 h from 1.17 ppm to 1.21 ppm. Increased consumption of isobutyric anhydride had not been observed in reactions

with propylamine. We hypothesised that this was a result of better peak separation, which underlines the importance of carefully selecting integral areas. However, it could also be a result of shorter timescales associated with reaction.

To test whether the apparent increase in anhydride consumption was a product of side reactions between isobutyric anhydride and water, chloroform-*d* was distilled over calcium hydride to remove adventitious moisture. N,N'-dibenzylethane-1,2-diamine was added to a solution of isobutyric anhydride in this chloroform-*d* in a Schlenk NMR tube, which was sealed to prevent atmospheric water from entering the sample over the course of the reaction. ^{1}H NMR analysis indicated that the course of the reaction was unchanged by the distillation protocol. Peak overlap was noted again in the first 3 hours, which was resolved as the doublet associated with isobutyric acid shifted downfield. The rates of reaction calculated by integration were comparable to those in un-distilled chloroform. On this basis, we hypothesise that the apparent increase in anhydride consumption could not be a product of reactions between the anhydride and water, as these rates would have been affected by the distillation procedure. Instead, we believe this is an artefact of the integration procedure, which could be improved using peak-fitting procedures to estimate the area under each curve. In further work, it would be useful to identify an anhydride with a chemical shift that does not overlap with amine peaks, although isobutyric acid was not found to affect integrations in reactions with reduced cages.

*3.2.4.2 Reactions between **RCC1** and acid anhydrides*

To test how conversions could be quantitatively measured in reactions between cage molecules and anhydrides, a slight stoichiometric excess (12.5 eq.) of isobutyric anhydride was added to a solution of **RCC1** in $CDCl_3$ (Figure 3.25a), and the reaction

was studied by ^{1}H NMR spectroscopy over 6 hours. Chloroform-*d* was distilled over calcium hydride, which removed any water present in the sample, and reduced the number of secondary hydrolysis reactions occurring.

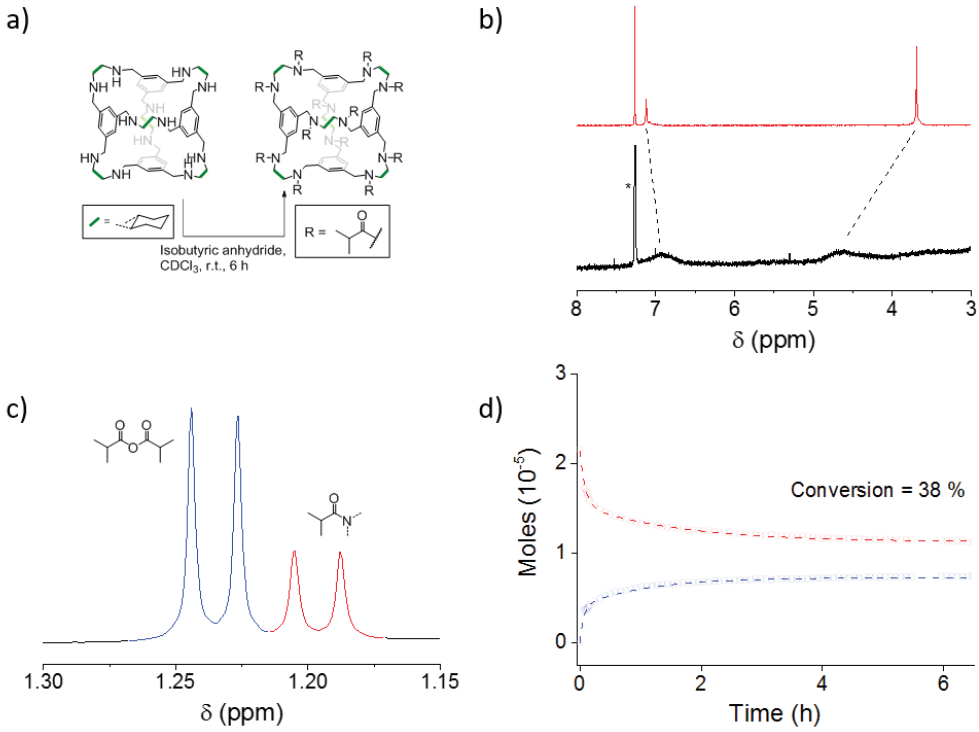


Figure 3.25. (a) Scheme for the reaction between RCC1 and isobutyric anhydride; (b) ^{1}H NMR spectra prior to addition of isobutyric anhydride (red) and after 6 hours of reaction (black). ^{1}H NMR spectrum focused from 8 to 3 ppm. Asterisked $CHCl_3$ peak has been truncated for clarity. Dotted lines connect chemical shifts that are structurally related between the reactants and products; (c) ^{1}H NMR spectrum of the crude reaction mixture after 6 hours, between 1.3 and 1.15 ppm. Highlighted, the chemical shift associated with isobtyric acid (blue) and isobutyric anhydride (red); (d) Quantities of isobutyric anhydride (red) and isobutyric acid (blue) in solution, measured relative to the internal capillary reference.

Conversions were calculated from the decrease in signal intensity at 1.19 ppm, associated with isobutyric anhydride, and the increase in signal intensity at 1.23 ppm, which was associated with isobutyric acid. Conversion was calculated to be 38 % (Figure 3.25d), below that of N,N'-dibenzylethane-1,2-diamine under the same

conditions. This corresponded to approximately 4.5 amine nitrogens per cage. The reaction rate was similar to that seen with N,N'-dibenzylethane-1,2-diamine. To confirm that reactant consumption was not partially a product of hydrolysis side reactions, the rates of anhydride-consumption and acid-production were compared (Figure 3.26). In this case, the rates were identical, which indicated reactions with water had not contributed to anhydride consumption.

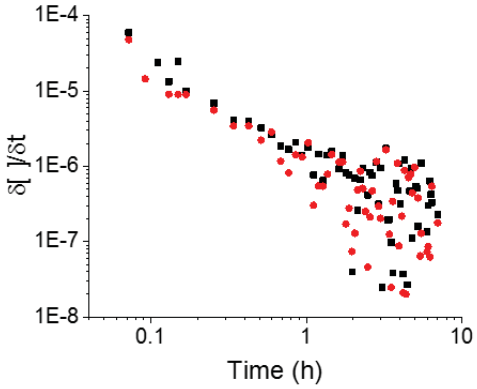


Figure 3.26. Double-log plot of reaction rates over time. Magnitude values are used.

As with **AA3**, a slight stoichiometric excess of anhydride was added to solutions of **RCC1** in $CDCl_3$. In each case, samples were studied in the presence of an internal capillary standard and average substitutions were calculated by comparing the concentration of isobutyric acid generated during the reaction with the known starting quantity of cage. Anhydrides were added by subsurface injection and the reaction was monitored by ^{1}H NMR spectroscopic analysis.

Between 5 and 15 μL of anhydride was added in each case, which changed the concentration of **RCC1** in solution by less than 1 %. Thus, changes in concentration brought about by the addition of anhydride starting materials were not considered in calculating the volumes of solution for ^{1}H NMR analysis.

In the case of reactions between **RCC1** and **AA1–3**, the initial rate at which **RCC1** was consumed were found to be almost independent of the anhydride steric bulk (**AA3**: Figure 3.25; **AA1**, **AA2**: Figure 3.27a–b). In the case of **AA4**, the reaction rate was noted to have decreased significantly (Figure 3.27c). Marked differences were observed between thermodynamic limitations on substitution. The average substitution after 6 hours was noted to decrease from 11.5 to 1 from **AA1** to **AA4**. Furthermore, thermodynamic limitations appeared to vary linearly with the number of carbons attached to Cα (Figure 3.27).

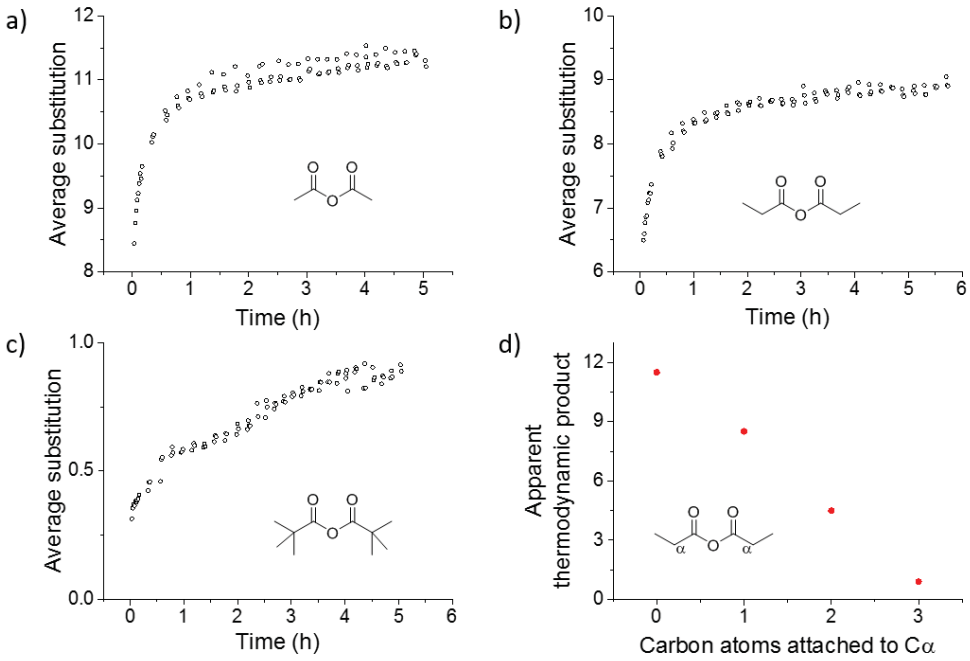


Figure 3.27: Average substitution of cages calculated by comparison between ^{1}H NMR integrals with internal capillary standard for reactions between RCC1 and (a) acetic anhydride, (b) propionic anhydride, and (c) trimethylacetic anhydride; (d) The apparent thermodynamic limit of substitution, calculated by the same method as in (a) for the point at which the reaction rate had decreased by 90 %.

3.2.4.3 Reactions between RCC3 and acid anhydrides

We hypothesised that, in addition to the effect of anhydride bulk, the steric environment surrounding amine groups within the cage would significantly affect the rate of reaction. To test this, a solution of **AA3** (31.4 μL, 0.021 mmol, 12 eq.) in chloroform-*d* (0.5 mL) was added to a solution of **RCC3** (20.0 mg, 0.0017 mmol, 1 eq.) in chloroform-*d* (0.5 mL) and the resultant reaction was monitored by ^{1}H NMR spectroscopic analysis for 6 hours.

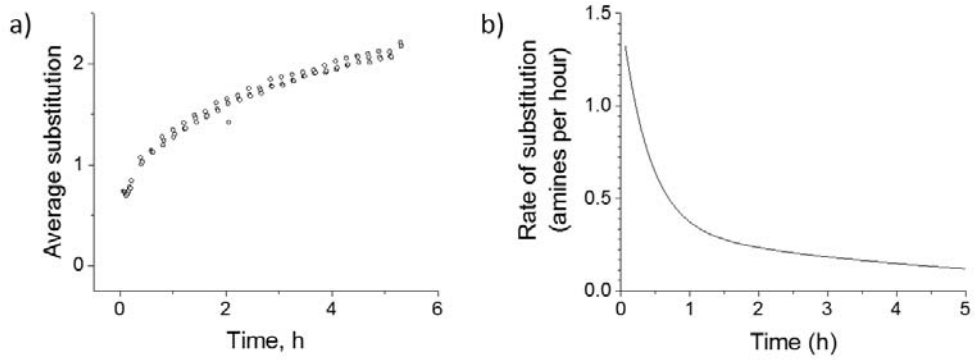


Figure 3.28: (a) Average substitution during a reaction between RCC3 molecules and AA3, measured by ^{1}H NMR spectroscopic analysis; (b) Rate of reaction between RCC3 and AA3.

The initial rate of reaction between **RCC3** and **AA3** was significantly lower than that associated with reaction between **RCC1** and **AA3** (Figure 3.27a compared to Figure 3.28). The rate of reaction was observed to decrease continuously over 5 hours, which was not consistent with a reaction in which a single substituent was favoured. However, we hypothesised that mixtures enriched in mono-acylated cage, **A_1RCC3,** could be made by adding substoichiometric quantities of **AA3**. This was predicted to generate a mixture of products weighted towards **RCC3** and **A_1RCC3**, and subsequently removing unreacted **RCC3**. Work by Liu et al. previously established a method by which the solubility of **RCC3** could be dramatically reduced by reaction

with either acetone or paraformaldehyde (Figure 3.29).[38] Thus, we hypothesised that reaction of crude mixtures with acetone or paraformaldehyde would increase the difference in solubility between **RCC3** and **A$_1$RCC3** and aid their separation.

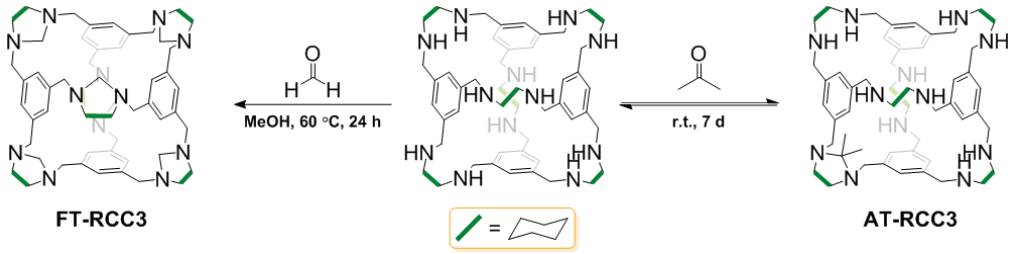


Figure 3.29. Reaction scheme depicting the reactions in which FT-RCC3 (left) and AT-RCC3 (right) are formed from RCC3.

3.2.5 Attempts to enrich crude cage reactions.

To test whether enriched mixtures could be generated, a substoichiometric quantity of **AA3** (55 μL, 0.035 mmol, 0.8 eq.) was added to a solution containing **RCC3** (500 mg, 0.044 mmol, 1 eq.), which analysed by ^{1}H NMR and mass spectroscopy. Mixtures of **RCC3** and **A$_1$RCC3** were confirmed by ^{1}H NMR spectroscopic analysis, which demonstrated characteristic sharp peaks associated with **RCC3**, in addition to the broadness associated with desymmetrised cage (Figure 3.30).

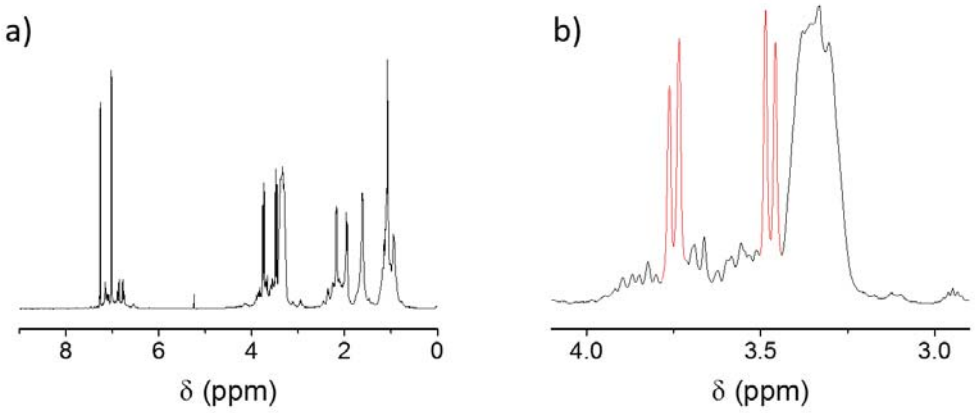


Figure 3.30. ^{1}H NMR spectrum of crude reaction mixtures between RCC3 and isobutyric anhydride; (a) full spectrum; (b) Spectrum focussed between 4 and 3 ppm, demonstrating chemical shifts produced by RCC3 (red) and by acylated cage (black)

Crude reaction mixtures were dried under reduced pressure, and a portion (100 mg) of the resultant solid was dissolved in acetone. After two days, a white precipitate had appeared on the inside of the glass. The supernatant was isolated by filtration and analysed by mass spectrometry, which indicated that solutions were enriched compared to crude mixtures (Figure 3.31). However, in spite of the ease with which solids could be isolated, the mass of material recovered from 100 mg of crude mixture was found vary between 35 and 45 mg. This was thought to be a result of the reversibility of the acetone-tying reaction, and the mass of material may be dependent on the quantity of water in solution.

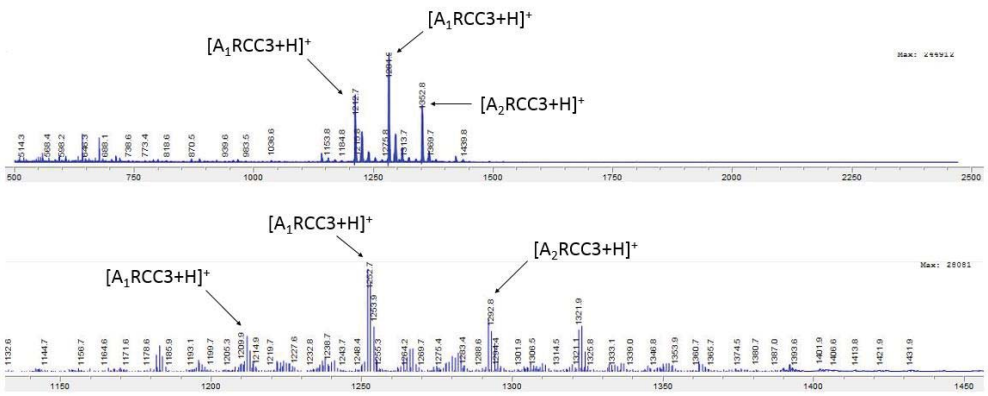


Figure 3.31. ESI mass spectra for: (a) the reaction between 0.8 equivalents of AA3 and RCC3; and (b) the supernatant following reaction with acetone, filtration and dissolution in a 8:1:1 mixture of chloroform, methanol and water.

We hypothesised that reactions between **RCC3** and paraformaldehyde, which were found to be irreversible,[38] may enhance the reliability of **RCC3** recoveries. To test this theory, three reactions were conducted in which **RCC3** (50.0 mg, 0.047 mmol, 1 eq.) and **AA3** (6.2 μL, 0.037 mmol, 0.8 eq.) were mixed for 4 hours. Solvents were removed under reduced pressure and the resultant solids were dissolved in a solution of paraformaldehyde (20 eq.) in methanol and heated to reflux for 16 hours. Solid material recovered by filtration was dried under vacuum, and found to be 30.7 ± 0.8 mg. ^{1}H NMR spectroscopic analysis of the solid confirmed it was FT-RCC3 (Figure 3.32, red spectrum), while analysis of the filtrate determined that the majority of RCC3 had been removed. Small quantities of **FT-RCC3** in the enriched sample were identified by chemical shifts at 4.02 and 3.23 ppm. A large number of peaks at 3.5 ppm in the crude ^{1}H NMR spectrum of the filtrate indicated that a number of partial reactions had occurred between paraformaldehyde and the acylated cage. We hypothesised that the presence of a substituted amide, which decreases the symmetry of the cage, may hinder reaction between amine groups and formaldehyde in solution.

Thus, a larger number of isomers were formed from the reaction between **A_1RCC3** and paraformaldehyde, than had previously been present in the mixture. Thus, on the basis of these reactions, it was determined that neither acetone nor paraformaldehyde could be used to enrich crude reaction mixtures that produced a mixture of acylated cages.

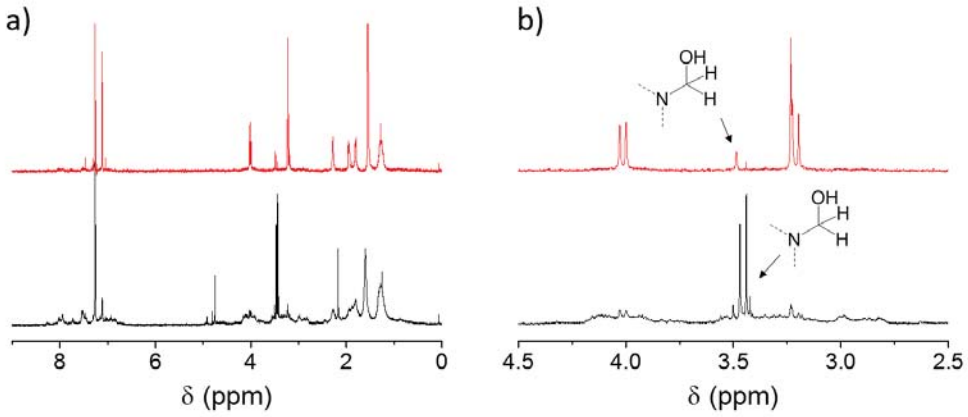


Figure 3.32. ^{1}H NMR specta of solid (red) and filtrate (black) material that was isolated on dissolving crude reaction mixtures between RCC3 and isobutyric anhydride in a solution of paraformaldehyde in methanol and heating to reflux; (a) full spectrum; (b) Spectrum focussed between 4.5 and 2.5 ppm.

However, we hypothesised that the ability to reproducibly recover **RCC3** from solutions would allow us to determine whether reactions between **RCC3** and **AA3** discriminated kinetically between low substitution states. This was achieved by modelling the reaction a two-stage process, in which the cage is substituted sequentially to afford **A_1RCC3** and **A_2RCC3** in serial reactions (Figure 3.33). Simultaneous reactions were not considered, because the sterically bulky environments of both **RCC3** and **AA3** were likely to make this unfavourable.

Figure 3.33. Kinetics of sequential reaction between RCC3 and an anhydride, A. The acid by-product has been omitted for clarity.

Because of the stability of the resultant amide, it was assumed that $k_1 >> k_{-1}$ and $k_2 >> k_{-2}$. Thus, where [**AA3**] << [**RCC3**], the concentration of **RCC3** at the end of the reaction can be used to estimate k_1 and k_2. To test this theory, a series of ten reactions was conducted in which 0.2 – 1.0 equivalents of **AA3** were added to **RCC3** in 0.1 equivalent increments and the masses of **FT-RCC3** recovered were used to calculate the quantity of **RCC3** that remained after reaction had completed (Figure 3.34). The recoveries of **RCC3** were compared to modelled data in which the values of k_1 and k_2 were changed until the root mean square (RMS) difference between calculated and experimental data was minimised. Modelled data indicated that $k_1 = 1.04\ k_2$, which confirmed that once **A_1RCC3** was formed in solution, it was likely to react with anhydrides at a competitive rate, which was likely to produce a significant quantity of **A_2RCC3** during reaction.

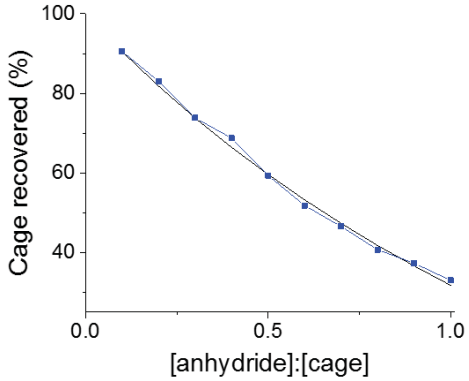


Figure 3.34. The mass of cage recovered (blue squares) with increasing quantities of anhydride. Data modelled for $k_1 = 1.02k_2$ (black line) is shown for comparison.

As chromatographic separation of **A_1RCC3** and more heavily-substituted products had found to be impossible due to the high polarity of amine nitrogens present, it was thought to be unlikely that pure **A_1RCC3** could be produced. To investigate why this level of reactivity may be possible, given the steric bulk associated with **RCC3**, a computational model produced by Kim Jelfs (ICL) indicated that cyclohexane groups within **RCC3** twisted considerably on reaction with an anhydride (Figure 3.35). Analysis of the computational model revealed that once one amine nitrogen had reacted, several more were directed externally. While this did not give a dynamic picture of behaviour in solution, it confirmed that it was possible for amine nitrogens within an acylated cage to react with additional anhydrides.

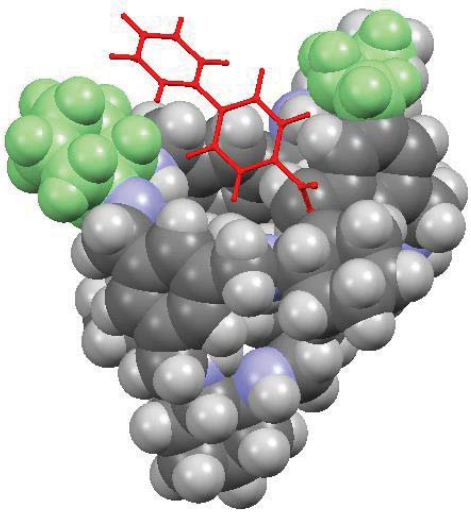


Figure 3.35. Computational model of the structure obtained through reaction of RCC3 with a single equivalent of a biphenyl anhydride. The structure of the anhydride has been highlighted in red and is shown as a stick representation. The remaining structure is space-filled, with cyclohexane groups adjacent to the biphenyl amide highlighted in green.

This degree of twist was thought to indicate that nitrogen atoms adjacent to cyclohexane rings within **RCC3** were able to react with anhydrides predominantly because the carbonyl carbon is bulky only on one side of the π-system. This allows the anhydride to point out into space, and places relatively little steric strain on the flexible structure of **RCC3**.

3.2.5.1 Synthesis of reduced desymmetrised cage $R1^{1}3^{5}$

As significant differences had been noted to exist between the initial rates of reaction with **AA3** and reduced cages, we hypothesised that a cage with two types of amine group may react differently at each site. To explore the possibility of regio-selective control *within* a cage, scrambled cages[20] were synthesised (see Section 3.1.3), which contained a mixture of cyclohexanediamine and ethylenediamine moieties (Figure 3.6).

As **RCC1** was found to react significantly more quickly than **RCC3** (see Figure 3.27 and Figure 3.28), it was thought that a cage with mixed ethylenediamine and cyclohexane diamine environments may react favourably in one environment when compared to another (Figure 3.36).

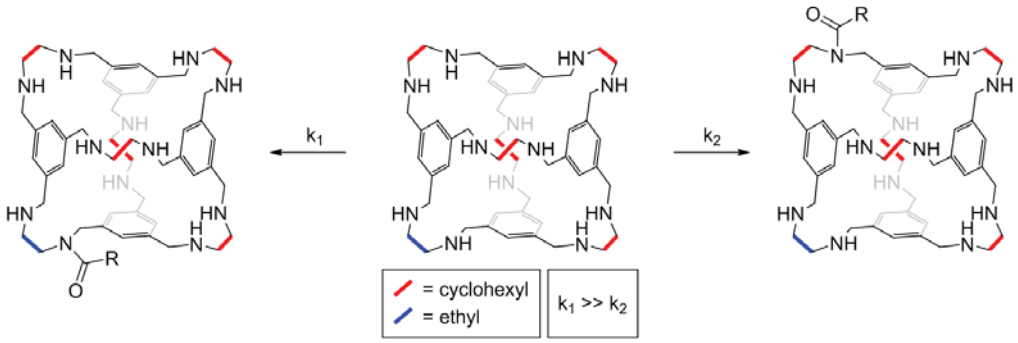


Figure 3.36. Scheme for the reaction between a cage with mixed amine environments and an anhydride. Reaction conditions and anhydride structure have been omitted for clarity.

To test this hypothesis, scrambled cages were synthesised as described in Section 3.1.3. The relative stoichiometries of ethylenediamine and cyclohexane diamine were varied to test the effect on the varying amounts of scrambled cage produced. As previously reported,[20] the naming convention used here is to label cages by vertex type: $\mathbf{1}^{x}\mathbf{3}^{y}$. Ethylenediamine, which is used in the production of **CC1**, is denoted **1**, and the number of ethylenediamine moieties in the cage is given by *x*. Similarly, cyclohexanediamine is denoted **3** and the number of cyclohexane diamine moieties in the cage is given by *y*. Crude reaction mixtures were analysed by HPLC, with peaks being assigned on the basis of their previously published elution order.[20] In order to maximise the effect of mixing ethylenediamine and cyclohexane diamine moieties, $\mathbf{1}^{1}\mathbf{3}^{5}$ was thought to present the best chance of favouring low cage substitutions. It was thought that in cases where $x > 1$, mono-substituted cages may compete for additional

substitution at similar rates as those where the first ethylenediamine moiety had already reacted.

It had previously been published that the most prevalent cage in the scrambled cage mixture generally the same diamine ratio as in the starting materials.[20] Thus, to favour the production of cages with a single ethylenediamine vertex (**$1^1 3^5$**) a 1:5 stoichiometric ratio of ethylenediamine to cyclohexanediamine was used during synthesis. This generated a crude reaction mixture in which **$1^1 3^5$** accounted for approximately 37 % of the sample (Figure 3.37a).

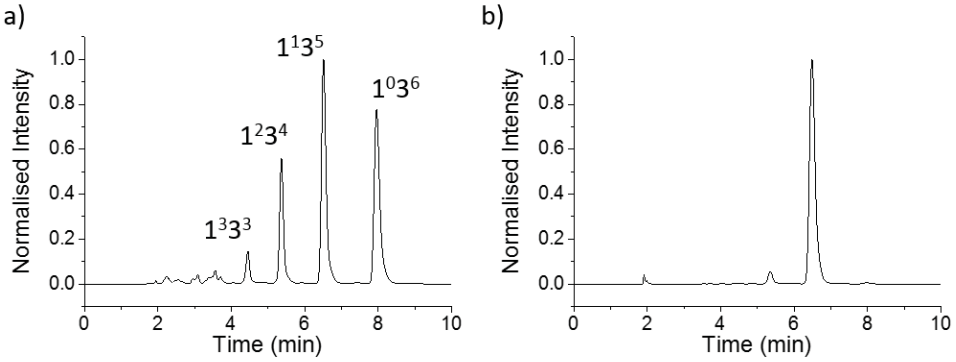


Figure 3.37. HPLC traces associated with (a) crude reaction mixtures and (b) purified samples of $1^1 3^5$.

Samples were purified by reverse phase chromatography. Because cage scrambling is a dynamic process, it had been reported that isolated **$1^1 3^5$** would equilibrate in solution to produce increasing quantities of **$1^2 3^4$** and **CC3**,[20] which was undesirable. In solution, **$1^1 3^5$** was found to be stable in $CHCl_3$, but equilibrated in methanol over 24 hours to a mixture of products, presumably due to the protic nature of the solvent. **$1^1 3^5$** was stable to equilibration in the solid state, although rigorous drying and air-tight storage was necessary to prolong life. Thus, the mobile phase was collected into a vessel cooled in an ice bath to slow equilibration, and solvents were removed quickly

under reduced pressure to afford a white solid enriched in **$1^1 3^5$**. HPLC analysis of the solid determined the purity to be 95 %area (Figure 3.37).

Samples of **$1^1 3^5$** were reduced to form **$R1^1 3^5$** using the modified procedure based on those used to reduce **CC1** and **CC3**. **$1^1 3^5$** (398 mg, 0.38 mmol, 1 eq.) was dissolved in chloroform, and diluted with methanol to form a dilute solution, to which was added $NaBH_4$ (600 mg, 15.7 mmol, 42 eq.) while maintaining a temperature of < 10 °C. This minimised equilibration both before and during reaction.

3.2.5.2 Reactions between $R1^1 3^5$ and acid anhydrides

To explore the steric control over reaction kinetics, the reactivity of **RCC3** and **$R1^1 3^5$** was tested in the presence of **AA3** at varying stoichiometries (Figure 3.38).

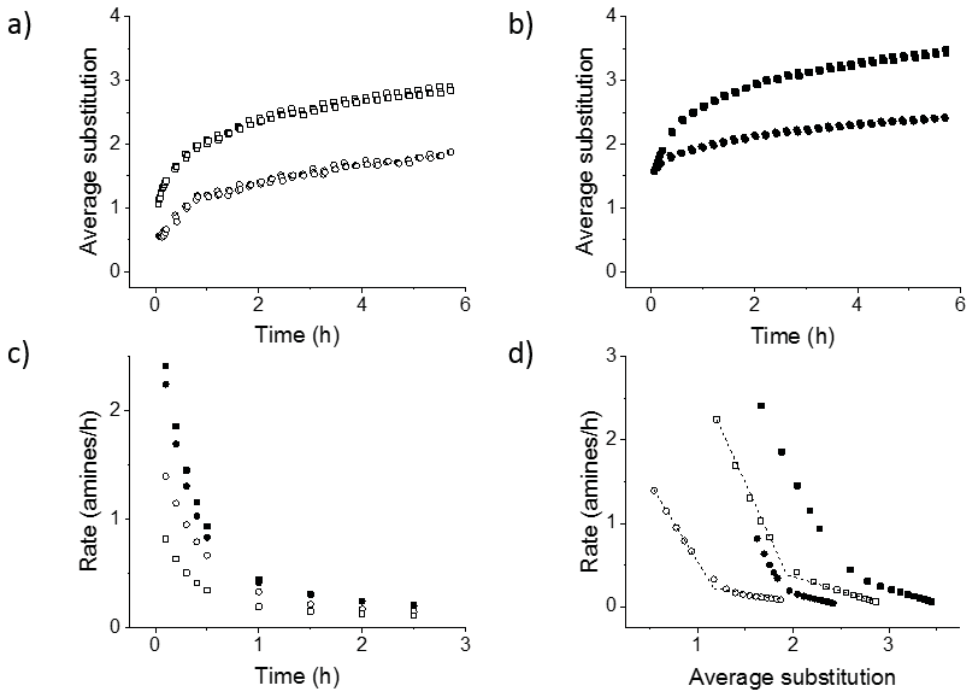


Figure 3.38. Average substitution of RCC3 (circles) and $R1^1 3^5$ (squares) in reactions with (a) 3 and (b) 6 equivalents of isobutyric anhydride; Rates of reaction, derived from the data in (a) and (b), plotted against (c) time and (d) average substitution.

R1^{1}3^5 was observed to react more quickly than **RCC3**. With both 3 (Figure 3.38a) and 6 (Figure 3.38b) equivalents of isobutyric anhydride, **R1^{1}3^5** had a consistently higher rate of reaction when compared with **RCC3**. To test whether the degree to which a cage was substituted had an effect on the rate of its reaction, rates were compared with the average substitution, as calculated by ^{1}H NMR integral area (Figure 3.38d). The rate at which average substitutions increase appeared to demonstrate a kinked shape, with a sudden decrease in rate at which substitution slowed. This occurred at 1 (**RCC3** with 3 eq.), 2 (**RCC3** with 6 eq. and **R1^{1}3^5** with 3 eq.) and 2.5 (**R1^{1}3^5** with 6 eq.) amines per cage (Figure 3.38). Increasing the relative concentration of isobutyric anhydride was observed to shift the position of change in kinetics from 1 to 2 amines per cage in the case of **RCC3** and 2 to 2.5 amines per cage in the case of **R1^{1}3^5**.

This two-stage decrease was inconsistent with the production of a single thermodynamically favourable product, which would be expected to result in a decrease in rate; however, an increase was instead observed. It was hypothesised that the upward shift in the sudden change of reaction kinetics from 1 to 2 amines per cage in the case of **RCC3** confirms that this is not a thermodynamic barrier. This is also true in the case of **R1^{1}3^5**, where the upward shift occurs from 2 to 2.5 amines per cage. As desymmetrisation had been noted to increase the solubility of cages we did not believe this was a result of differences in solubility between components.

In the case of **R1^{1}3^5**, mass spectroscopic analysis demonstrated species at 1211.8 and 1281.7 m/z, which confirmed the presence of mono- and di-acylated cages. However, determining whether substitution had occurred on amine nitrogen atoms adjacent to ethylenediamine groups was hampered by broadening in the ^{1}H NMR spectrum (Figure 3.39). Chemical shifts at 3.4 ppm in the ^{1}H NMR spectrum of R1^{1}3^5

were associated with C*H*$_2$Ar protons adjacent to ethylenediamine nitrogen atoms. Chemical shifts between 3.5 and 4.1 ppm were associated with CH_2Ar protons adjacent to cyclohexane groups (Figure 3.39). However, ^{1}H NMR spectroscopic analysis of material that had an average substitution of 1 amide group per cage demonstrated significant broadening, in spite of the expectation that this substitution would occur predominantly at the ethylenediamine nitrogen groups.

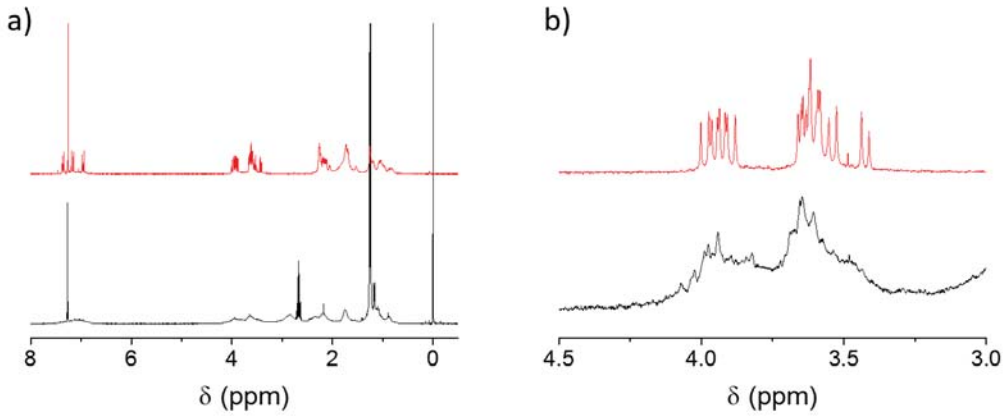
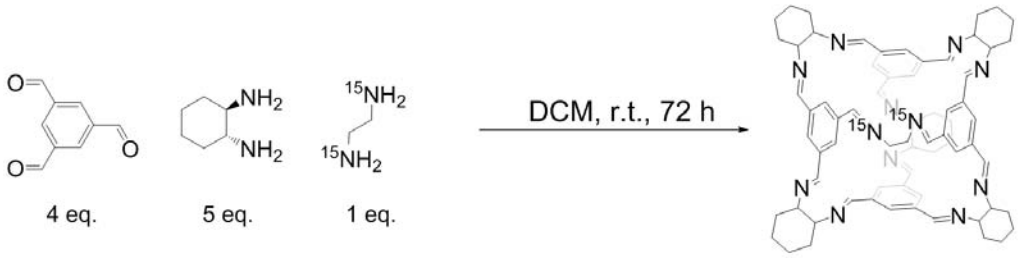


Figure 3.39. ^{1}H NMR specta of R1^{1}3^5 (red) and a crude mixture for the reaction between R1^{1}3^5 (black). (a) Full spectrum; (b) Spectrum focussed between 4.5 and 3.0 ppm.

Thus, in order to analytically resolve which species were created, it was necessary to synthesise an enriched isomer of **R1^{1}3^5** in which the ethylenediamine nitrogen atoms were ^{15}N, rather than ^{14}N. As ^{15}N has a spin of ½, ^{15}N NMR spectroscopy could be used to determine whether these nitrogens were substituted after reaction with **AA3**.

3.2.5.3 Synthesis and reactivity of enantiomerically enriched R1^{1}3^5

The scrambled cage in which the ethylenediamine moiety is enantiomerically enriched is termed **{^{15}N}1^{1}3^5**. This was synthesised using ^{15}N-isotopically labelled ethylene diamine in a procedure identical to that used to synthesise **1^{1}3^5** (Scheme 3.1).



Scheme 3.1. The synthesis of R1^{1}3^{5} by condensation reaction using nitrogen-labelled ethylene diamine.

^{15}N NMR spectroscopic analysis demonstrated one chemical shift at 326.2 ppm, and a second chemical shift at 326.9 ppm, which corresponded to 4 % by integral area (Figure 3.40a). This confirmed that the nitrogen atoms in **{^{15}N}1^{1}3^{5}** were chemically and magnetically identical. The second chemical shift at 326.9 ppm suggested that **{^{15}N}1^{2}3^{4}**, which remained present at a concentration of ~4 % (a/a) by HPLC analysis, could be distinguished by ^{15}N NMR spectroscopy.

{^{15}N}1^{1}3^{5} was reduced under the same conditions as had been used previously with **1^{1}3^{5}**. After synthesis, ^{15}N NMR spectroscopic analysis demonstrated a single chemical shift at 38.0 ppm, consistent with a single ^{15}N environment. The absence of a second chemical shift (Figure 3.40b) suggested that unsubstituted amine nitrogen atoms could not be differentiated in the same way as the imine nitrogen atoms.

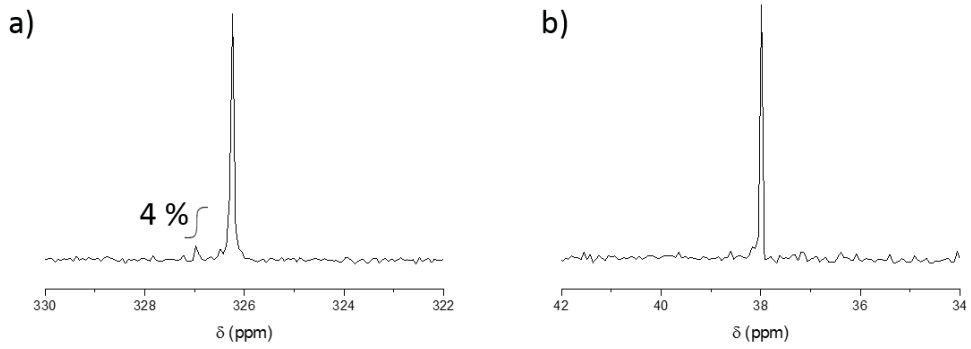


Figure 3.40. ^{15}N NMR spectra of (a) {^{15}N}1^{1}3^{5} and (b) R{^{15}N}1^{1}3^{5}.

Mass spectroscopic analysis demonstrated a single mass of 1089.8367 m/z (Expected: 1089.8364), which was consistent with a single, fully reduced cage. Evidence of **R{^{15}N}1^{2}3^{4}** was undetectable by MS. Because of its high polarity, analysis by HPLC was found to be impossible, as it eluted at the solvent front in a variety of solvents, and with a variety of column types.

Therefore, to determine which nitrogen atoms were more reactive, 2 equivalents of isobutyric anhydride were added to **R{^{15}N}1^{1}3^{5}** and the reaction was analysed by ^{1}H NMR spectroscopy. The organic solvents were removed under reduced pressure and the resultant solids were dissolved in chloroform-*d*. ^{1}H NMR spectroscopic analysis confirmed quantitative conversion of isobutyric anhydride to isobutyric acid, indicating that two equivalents had reacted with the isotopically enriched cage.

To confirm the location at which isobutyric anhydride had reacted with the reduced cage, ^{15}N NMR spectra were collected. ^{15}N NMR spectroscopic analysis demonstrated partial disappearance of the chemical shift at 38.0 ppm (Figure 3.41). However, the spectrum also demonstrated the presence of six acylated species.

Chemical shifts associated with addition products were noted to take the form of doublets, which were generated ^{2}J coupling to the NCH_2 protons on the ethylenediamine backbone.[39] Nitrogen atoms were distinguishable from each other, which resulted from the chirality of the cage structure,[15] generating pairs of doublets (Figure 3.41b–c). The magnitude of diastereomeric splitting was noted to be significantly larger in the case of one doublet pair (Figure 3.41b, red), which was consistent with the desired product in which both ^{15}N nitrogen atoms were substituted.

The large downfield chemical shift change from 38 to 68 ppm was associated with conversion from amine to amide nitrogen atoms.[40] Thus, several other doublet pairs in this region were assigned to cages in which the desired product had proceeded to react

further (Figure 3.41a). Chemical shifts between 34 and 40 ppm were thought to be associated with reaction products in which **R{^{15}N}1^{1}3^{5}** reacted at non-enriched sites.

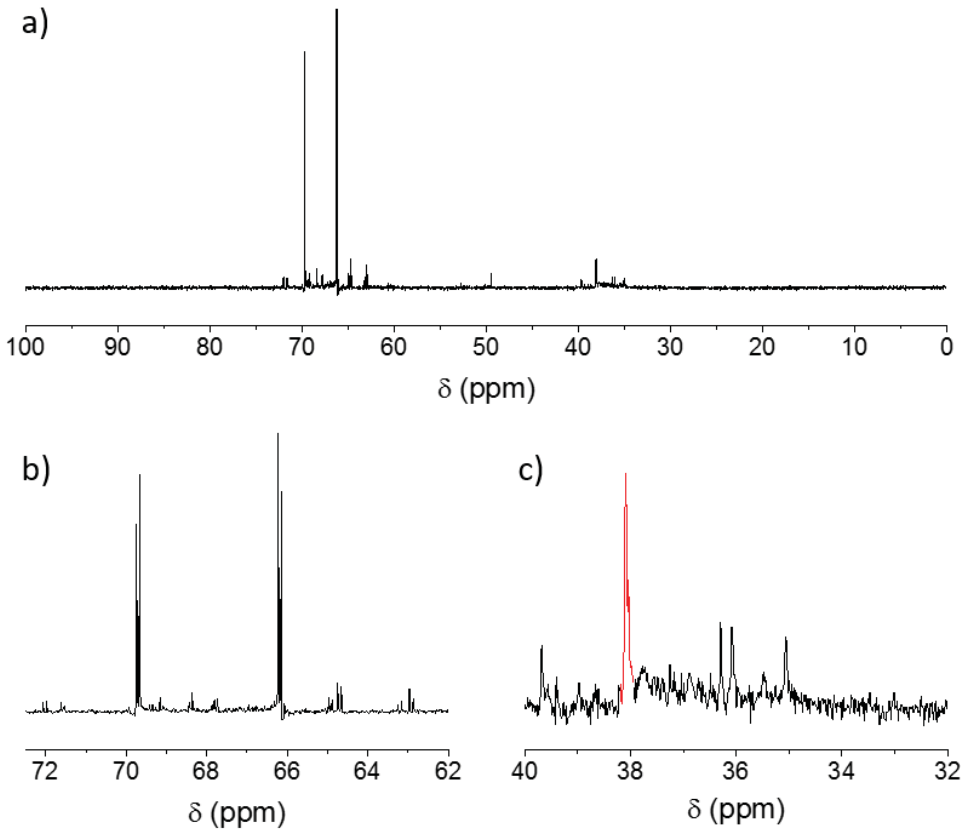


Figure 3.41. ^{15}N NMR spectrum of a crude reaction mixture between R{^{15}N}1^{1}3^{5} and isobutyric anhydride (2 eq.); (a) Full spectrum; (b) chemical shifts associated with products from 72 to 62 ppm; (c) chemical shifts associated with starting material, and products from 40 to 32 ppm.

Peaks associated with the major product accounted for 73 % of the spectrum by integral area. However, attempts to separate these materials by either recrystallization or chromatography were unsuccessful. Substitution of **R{^{15}N}1^{1}3^{5}** appeared to make cages more soluble, although differences in solubility were insufficient to differentiate between compounds in common organic solvents. Several attempts were made to purify material by dissolving crude mixtures in DCM and precipitating material

through the addition of an anti-solvent such as hexane, pentane or toluene. However, in all cases, the solid material isolated was no purer than the crude material.

Products contained large numbers of amine nitrogen atoms, which made separation by reverse-phase chromatography difficult. All compounds eluted at the solvent front in aqueous and mixed aqueous-glycol solutions with a variety of stationary phases. Using normal phase chromatography, the reduced cages remained adhered to both silica and amino-functionalised silica even with highly polar solvent mixtures such as methanol/DCM.

A final attempt to enrich samples of impure acylated **R{^{15}N}1^{1}3^{5}** was made by dissolving synthesised material in acetone, and subsequently by reacting crude materials in a solution of paraformaldehyde in methanol. This was based on previous work, which had identified two reactions that could be performed on **RCC3**, that rendered material insoluble (Figure 3.29).[38]

In the case of **AT-RCC3**, **RCC3** was found to become insoluble in acetone through the reversible formation of an aminal adduct. This presented the opportunity to recover unreacted **R1^{1}3^{5}**, and simultaneously enrich crude reaction mixtures. To test the effectiveness of this first technique, crude reaction mixtures were dissolved in acetone and the resultant clear solution was allowed to stand for 10 days.

No solid precipitate was recovered, and the volume of solution was reduced by nine tenths. On subsequent standing, a solid white precipitate was produced, which was insoluble in common organic solvents. This may be a polymeric adduct, in which several cages have become ‘tied’ with paraformaldehyde, as a result of the outward-facing ethylenediamine nitrogen atoms. In spite of this, the remaining supernatant was found to be unenriched. Finally, crude reaction mixtures were dissolved in a solution of paraformaldehyde in methanol and the resultant clear solution was heated to reflux

for 16 hours. Analysis by ^{1}H NMR spectroscopy demonstrated a significant increase in peak broadness, which made structural analysis impossible (Figure 3.42a). This indicated that a large number of reaction products had formed, and suggested that difficulties in tying acylated **R1^{1}3^{5}** with paraformaldehyde exceeded those experienced in the tying of acylated **RCC3**. To confirm that peak broadness was not indicative of paraformaldehyde-tying in desymmetrised cages more generally, **R1^{1}3^{5}** was dissolved in a solution of paraformaldehyde in methanol, and heated to reflux for 16 hours. After this time, ^{1}H NMR spectroscopic analysis demonstrated a single component, with a small number of chemical shifts (Figure 3.42b). This confirmed that peak broadness observed in ^{1}H NMR spectra of reactions between acylated **R1^{1}3^{5}** and paraformaldehyde were most likely the result of a large number of isomers being formed from partial reaction completion.

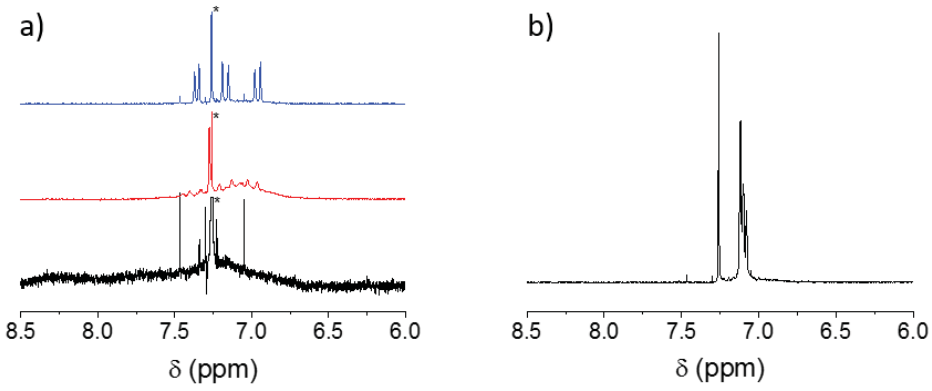


Figure 3.42. ^{1}H NMR spectra of R1^{1}3^{5} (blue), a crude mixture for the reaction between R1^{1}3^{5} and AA3 (red), and a crude mixture for the reaction between acylated R1^{1}3^{5} and paraformaldehyde (black). Asterisked peaks correspond to $CHCl_3$, which have been truncated for clarity.

3.2.6 Reactions between reduced cages and symmetrical ketones.

We hypothesised that the high degree of reactivity demonstrated by nitrogen atoms adjacent to cyclohexane groups may be a result of conformational flexibility (see Figure 3.35). Thus, we hypothesised that cyclohexyl moieties adjacent to nitrogen atoms in **RCC3** may prohibit reaction with a symmetrical ketone, where reaction with an unsymmetrical anhydride had been possible. Conversely, we hypothesised that nitrogen atoms adjacent to the ethylenediamine vertices would not be sterically restricted. To test this, a library of ketones was selected to probe whether it might be possible to thermodynamically differentiate between ethyl and cyclohexyl moieties within a scrambled cage (Table 3.2).

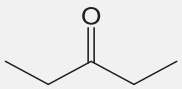
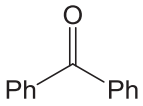
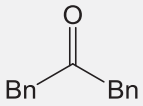
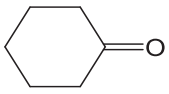
Entry	Ketone	Abbrev.
1		**SK1**
2		**SK2**
3		**SK3**
4		**SK4**

Table 3.2. Commercially available symmetrical ketones used in this study.

To test the reactivity of symmetrical ketones with ethylenediamine cage vertices, 2.1 mg (0.25 mmol) of **SK1** and 0.5 mg (0.024 mmol, 0.1 eq.) of *p*-toluenesulfonic

acid (PTSA) were added to a solution of RCC1 (200 mg, 0.25 mmol, 1 eq.) in d_4-methanol and the resultant mixture was analysed by ^{1}H NMR for 5 days. PTSA had previously been demonstrated generally to act as an acid catalyst,[41,42] and was added to aid reaction between the amine moieties in **RCC1** with the carbonyl carbon in **SK1**.

Initial screening by ^{1}H NMR spectroscopic analysis indicated ketone consumption was complete after 5 days. In the case of pentan-3-one, **SK1,** this was identified by the disappearance of the triplet chemical shift and appearance of a quartet at 1.0 ppm. The quartet at 2.45 ppm was observed to disappear completely (Figure 3.43).

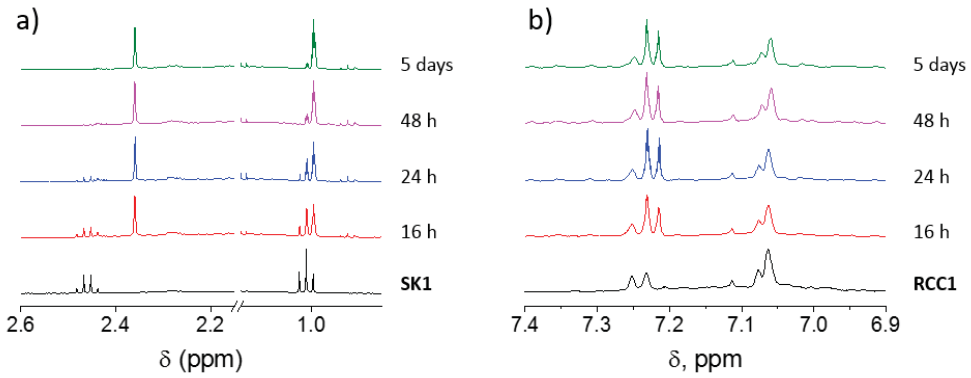


Figure 3.43. ^{1}H NMR spectra of crude reaction mixtures between RCC1 and SK1 focussed around chemical shifts associated with (a) SK1 and (b) RCC1.

In spite of some conversion being evident, the multiplet chemical shift associated with **RCC1** at 7.07 ppm was observed to remain unchanged even after 5 days (Figure 3.43). In the synthesis of **AT-RCC3**, a significant change was observed in the chemical shifts associated with aromatic protons, due to desymmetrisation of the aromatic protons.[38] This indicated that bonding at the amine nitrogen could significantly affect the electronic environment around aromatic protons. However, in spite of conversion noted in **SK1**, the aromatic protons associated with **RCC1** were

not observed to change over 5 days. Mass spectroscopic analysis of crude reaction mixtures confirmed that the *m/z* associated with **RCC1** (817.5 Da) was retained throughout the reaction, and no masses associated with addition products were noted. This appeared to indicate that **RCC1** remained unreacted across the course of the reaction. Furthermore, no change was observed in the chemical shifts associated with PTSA.

This pattern was repeated in **SK2–4.** In each case, conversion was noted in the chemical shifts associated with the ketone, whilst chemical shifts associated with PTSA and cage remained unchanged.

To investigate whether an adduct had been formed, ^{1}H PFG-NMR analysis was undertaken on each sample (Table 3.3). In each case, the diffusion coefficient of the ketone was observed to remain constant throughout the reaction, and the diffusion coefficient of the ketone product was identical to that of the ketone starting material.

Entry	Ketone	Amine	D (×10-10 m2 s-1)		
			Amine	PTSA	Ketone
1	SK1	R1135	4.45	7.30	10.7
2	SK2	RCC1	2.56	6.91	9.93
3	SK3	R1135	4.60	9.07	20.1
4	SK4	RCC1	5.09	9.07	15.3

Table 3.3. Diffusion coefficients measured from crude reaction mixtures between a series of ketones and cage materials.

With the exception of the reaction between **EDA** and **SK1**, significant differences remained between the diffusion coefficients of amine and ketone chemical shifts even

after reaction completion. This indicated that the materials diffused through solution at different rates, which suggested no chemical bonds had been formed between the species. Instead, systematic proton exchange may have occurred between ketone and solvent, facilitated by the presence of both PTSA and amine functional groups (Figure 3.44).

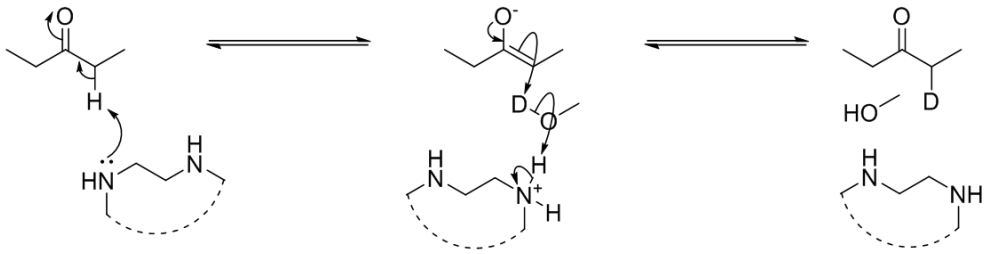


Figure 3.44. Reaction scheme for proton exchange between deuterated solvent and SK1. Tautomerisation of enolate to ketone *via* proton exchange with methanol-*d* and cage is drawn as a single step for simplicity.

To investigate whether this mechanism was correct, reactions were conducted in the absence of PTSA, and subsequently in the presence of PTSA, but absence of **RCC1**. This was supported by the absence of reaction in cases where PTSA was not present. Furthermore, no reaction was noted in cases where PTSA was present, but the amine was excluded.

3.2.7 Understanding host-guest chemistry in a Porous Liquid

Although it was impossible to purify cage monomers, one significant benefit of the scrambled cages synthesised in the previous Section is their increased solubility. This is a product of their decreased solid-state interactions (see Section 3.1.3). It was previously hypothesised that it may be possible to dissolve a porous molecule in a

solvent that was too large to enter the pores.[43] This would produce a mixture in which gas could be trapped inside the porous molecule that would not be ejected as a result of competitive interactions with solvent. This mixture was termed a "Porous Liquid." The first porous liquid, **PL1**, described below, was limited by the low solubility of the host molecule. In the following Section, we use scrambled cages to generate a porous liquid with a higher capacity for gas uptake.

3.2.7.1 Synthesis of a novel porous liquid

In the absence of a kinetic barrier that prevents solvent entering the host, high host-gas binding strengths are difficult to achieve. However, host-guest binding can be strengthened by making solvent-host interactions repulsive. Examples of this include hydrophobic binding in aqueous solvents and ion-ion interactions in non-polar solvents.[21] In the case of porous liquids, a binding pocket is produced that is kinetically- or thermodynamically-inaccessible to solvent. Using this strategy, it is possible to achieve much higher binding strengths even for poor guests (Figure 3.45).

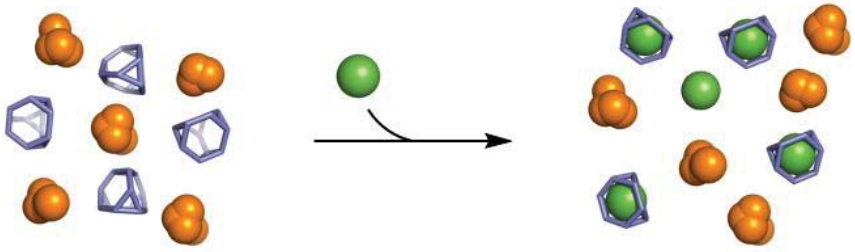


Figure 3.45. Schematic representations of two types of host-guest systems; (a) thermodynamically-unfavourable solvent-host interaction favours guest-uptake, pushing equilibrium to the right; (b) kinetically-inaccessible molecular sites preclude solvent-host uptake. Addition of a guest results in immediate uptake. "Empty" cages are likely filled with gas, which has been omitted for clarity. Solvent, guest and bulky solvent molecules are shown as red, green and orange respectively. Cages are represented as blue truncated tetrahedra.

The first example of a porous liquid (**PL1**) was developed by Cooper and James,[44] and involved the synthesis of cages with large external crown-ether components, which increased the solubility of these cages in 15-crown-5. The ether solvent, unable to penetrate the cage window because of its size, rendered the cavities of the cages empty (gas-filled) in the absence of specific guests (Figure 3.46).

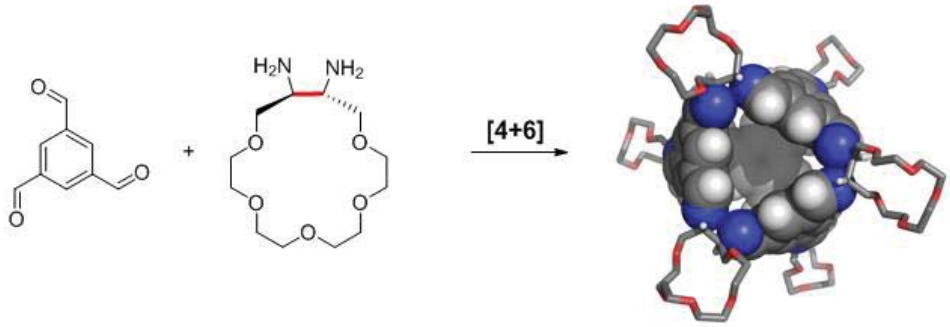


Figure 3.46. Synthesis of the first Type II porous liquid, a crown-ether cage, which contains cages with cavities that are kinetically-inaccessible to solvent.

In the solid state, porous cage materials mostly have moderate surface areas (< 1000 m^2 g^{-1}),[15] although these are constantly being improved.[45–47] However, by diluting cages into a non-porous medium, the pore density decreases as the volume occupied by solvent accounts for the majority of the material.

The high mass of the crown-ether cage (2102 Da), coupled with finite solubility was found to limit potential capacity, which would impact practical applications. A cage to solvent ratio of 1:12 was found to correspond to a void fraction of just 0.7 %.[44] The void fraction represents the volume within the material that is permanently porous, and is an indication of the capacity of the medium for storage.

To optimise capacity, either pore volume or solubility can be increased. In order to increase the pore volume, significant redesign of cage structure would be required as this is a property that is dependent on the size and geometry of the starting materials.

Increasing solubility, which can be accomplished by reducing solid-state cage-cage interactions and through solvent screening, was thought to be less labour-intensive. One technique for decreasing solid state interactions of organic cage molecules is desymmetrisation through dynamic covalent scrambling[20].

Using this theory, Becky Greenaway (UoL) has recently developed a second porous liquid, **PL2**,[44] with significant increases in solubility over **PL1**. This was accomplished by systematically screening desymmetrised cage mixtures for solubility in a variety of solvents.

In spite of their high symmetry, tetrahedral cages were chosen as the base topology for the porous liquid. Although there is some evidence of tetrahedral molecules that are prone to structural collapse,[48] the majority of T_h cages are structurally persistent.[45–47,49] This ensures that molecules retain cavities that are capable of adsorbing guests.[17,50]

As described in Section 3.1.3, the introduction of ethylenediamine into the synthesis of **CC3** produces scrambled cage mixtures with a significantly higher solubility than the parent molecules, **CC1 and CC3**. To further increase the solubility of cage mixtures, Becky Greenaway (UoL) introduced a desymmetrised diamine (1,2-diamino-2-methylpropane, DMP) in place of 1,2-ethylenediamine.[44] Thus, in addition to the positional desymmetrisation caused by random placement of cyclohexane and ethylene moieties, this produced regiochemical desymmetrisation, in which the orientation of each diamine was also randomised (Figure 3.47). Each positional isomer gives rise to 2^n regioisomers, where n represents the number of diamines in the molecule. This way, the number of isomers was increased to 157, compared to the original **CC1**/**CC3** scrambling, which produced 10. As DMP has been used previously

in the synthesis of a tetrahedral cage, **CC13**,[51] the scrambled cage produced is referred to as $\mathbf{3^x13^y}$, where x and y refer to the number of amine molecules in the cage.

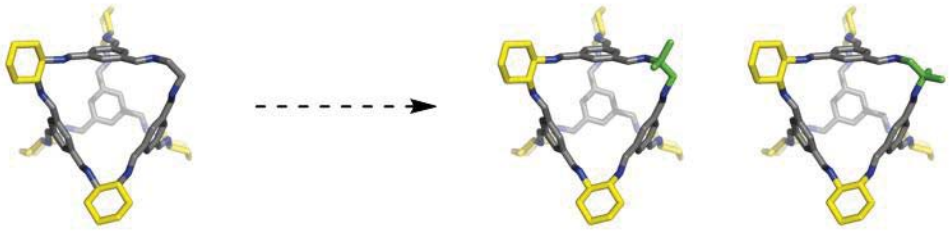


Figure 3.47. Desymmetrisation of one diamine component added complexity to the synthesis by producing regioisomers. The chemical structures of the regioisomers generated by DMP are shown with methyl groups highlighted in green.

One important factor in generating a porous liquid is kinetically preventing solvent molecules from entering the cage cavity. In this way, gas molecules in the cavity prior to dissolution are not ejected by competitive interactions from solvent molecules. To ensure that solvents screened were kinetically barred from cages, computational molecular dynamics (MD) were conducted by Daniel Holden (UoL), which determined whether selected solvent molecules were kinetically able to diffuse within the internal cavity of the scrambled cage molecules.

To determine the kinetic accessibility of cage cavities to solvent molecules, the Amorphous Cell module of Materials Studio was used to generate a scrambled cage solution containing 30 $\mathbf{3^313^3}$ molecules. Only one isomer was used, and it was assumed that this would not affect the results. DL_FIELD[52] was used to solvate the system with the chosen solvent. MD simulations were subsequently carried out using DL_POLY_2.20.[53,54] Further computational details can be found in Section 4 of the Supplementary Information of Reference 55.

Using solvents that had been identified as being too large to enter the cavity of the cage, Becky Greenaway screened scrambled cages and solvents for solubility and cage

stability.[55] This revealed that a scrambled cage mixture was most soluble in hexachloropropene (HCP). In traditional solvents, such as chloroform or water, the pores of cage molecules are filled with solvent. However, because the HCP solvent molecules were unable to enter through the window of the cage, cage pores, which are filled with air in the solid state, remained filled with air even in solution. The maximum solubility of scrambled cages in HCP was determined to be 232 mg mL^{-1}.[44] However, to ensure consistency throughout measurements, saturated solutions were not used.

To synthesise the scrambled cage used in the porous liquid, a solution of 1,2-diamino propane hydrochloride (204 mg, 1.39 mmol, 3 eq.) and trimethylamine (140 mg, 1.39 mmol, 3 eq.) in DCM (9 mL) and a solution of (R,R)-1,2-diaminocyclohexane (158 mg, 1.39 mmol, 3 eq.) in DCM (9 mL) were added to a solution of benzene-1,3,5-tricarboxaldehyde (300 mg, 1.85 mmol, 4 eq.) in DCM (90 mL) and the resultant solution was stirred for 3 days. After concentration *in vacuo*, the resultant solid was extracted into THF (50 mL) and the insoluble trimethylamine-hydrochloride salts were removed by filtration. After drying in vacuo, the scrambled cage was dissolved in HCP at a concentration of 200 mg mL^{-1} to afford **PL2**.

3.2.7.2 Determining gas uptake by NMR

To assess whether pores could permanently store gas molecules, CH_4 was introduced into the system by bubbling pure CH_4 gas through **PL2**. Saturation tests conducted by Becky Greenaway (UoL) confirmed that saturation of **PL2** was achieved after a flow of 60 mL of CH_4 per minute per mL of porous liquid. Initial ^{1}H NMR spectroscopic analysis revealed a significant shift in the ^{1}H NMR peak associated with CH_4 when compared to its chemical shift in neat HCP (Figure 3.48).

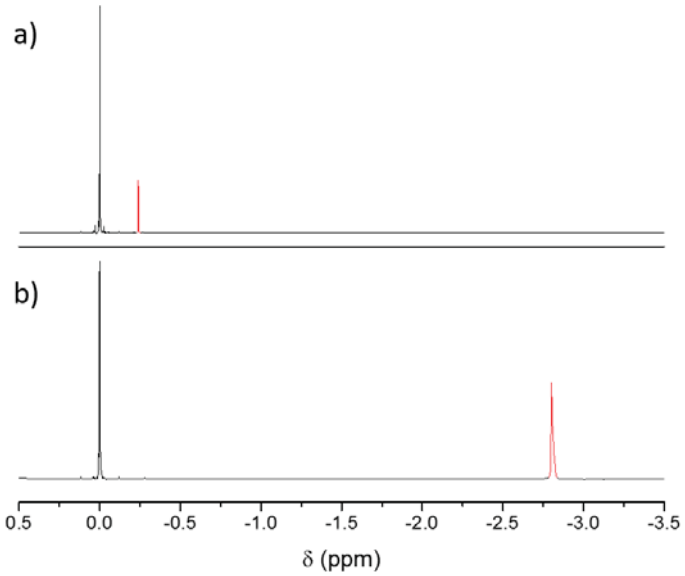


Figure 3.48. ^{1}H NMR spectra of CH_4 in (a) neat HCP (b) PL2. Peak is associated with CH_4 has been highlighted in red.

This shift, from -0.23 ppm, to -2.83 ppm, has been reported to indicate uptake of methane within the cavity of a molecule.[56,57] The upfield shift is thought to be related to the dense electronic environment within the cavity, which shields it from the effects of the permanent magnet. Thus, the magnitude of the shift is thought to be related to the electron density surrounding the methane, and thus the strength with which the methane is bound.

To rule out the possibility of a non-cavity-related cause for chemical shift change, such as imine-coordination, a structurally-analogous molecule with no cavity was synthesised by Becky Greenaway (UoL). This was designed to resemble the scrambled cage so that imine-CH_4 interactions would be similar, but did not contain a persistent void (Figure 3.49).

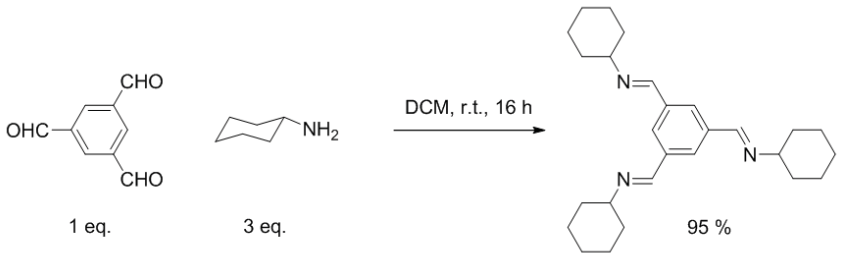


Figure 3.49. A non-cage analogue was synthesised by Becky Greenaway.

This small imine control (SIC) was intended to be tested alongside the porous liquid to assess both the chemical shift change associated with doped CH_4 and the apparent size of CH_4 in solution. The chemical shift change to a more high-field position has been attributed to guest incorporation in the higher electron density environment of a host molecule.[56,57]

Diffusion through solution is mostly dictated by molecular size. Thus, where one molecule is bound strongly to a second, they both diffuse at the same rate.[58–61] Thus, to confirm that this was the case, ^{1}H PFG-NMR spectroscopy was conducted on **PL2**, as well as a solution of SIC (200 mg mL^{-1}) in HCP, and neat HCP itself. Each sample was dosed with CH_4 identically, as described above and measured by ^{1}H PFG-NMR. Measurements were calibrated individually for each component in order to determine their size (Figure 3.50).

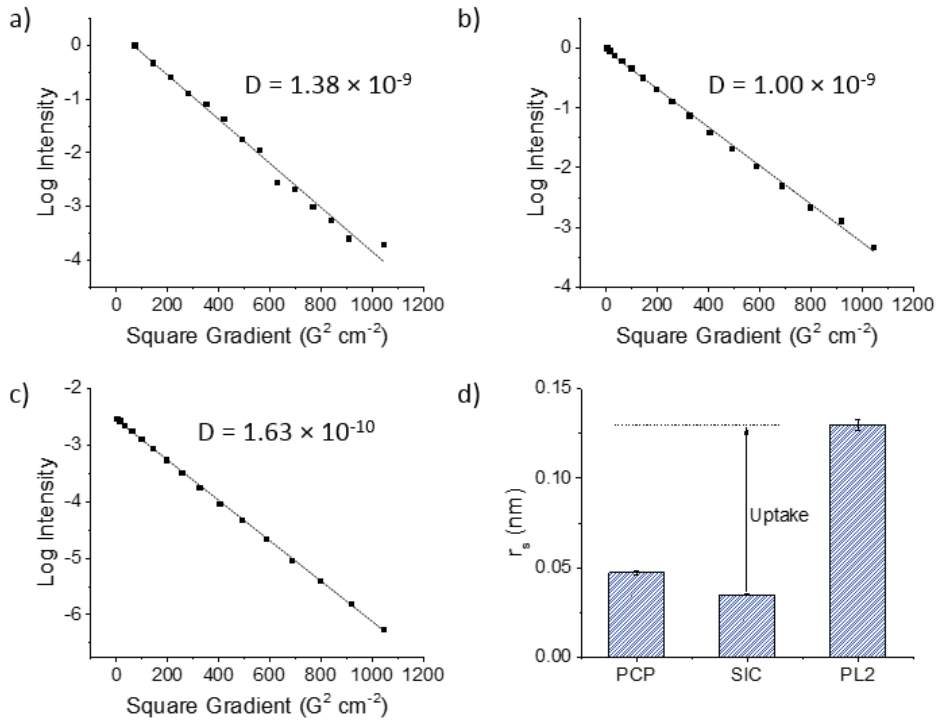


Figure 3.50. ST-plots of ^{1}H PFG-NMR for small imine control in (a) HCP, (b) a solution of the SIC in HCP at 200 mg mL^{-1}, and (c) PL2 at 200 mg mL^{-1}. In each case, diffusion coefficients extracted from data are shown; (d) Solvodynamic radii (r_s) of solutions in (a–c) calculated from the average of three separate experiments. Error bars represent 1σ. Increase in relative size on host-binding is highlighted by a vertical arrow marked "uptake."

^{1}H PFG-NMR spectroscopic analysis demonstrated that the apparent size of CH_4 in **PL2** was greater than in both the neat HCP and SIC solution. This confirmed that uptake was the result of cage-specific behaviour. At this time, it was unknown whether CH_4 uptake was the result of CH_4 molecules being adsorbed into single cage cavities. At such high concentrations, it was hypothesised that larger cage aggregates may have formed, which may also have permanent extrinsic pores between cage molecules.

To determine whether cages aggregated at such high concentrations, samples of **PL2** were prepared at concentrations of 25, 50, 100, 150 and 200 mg mL^{-1}. Because the viscosity of samples changed over this region, viscosities were measured rather

than using the native viscosity of the solvent (Figure 3.51). Viscosities were measured by Becky Greenaway.

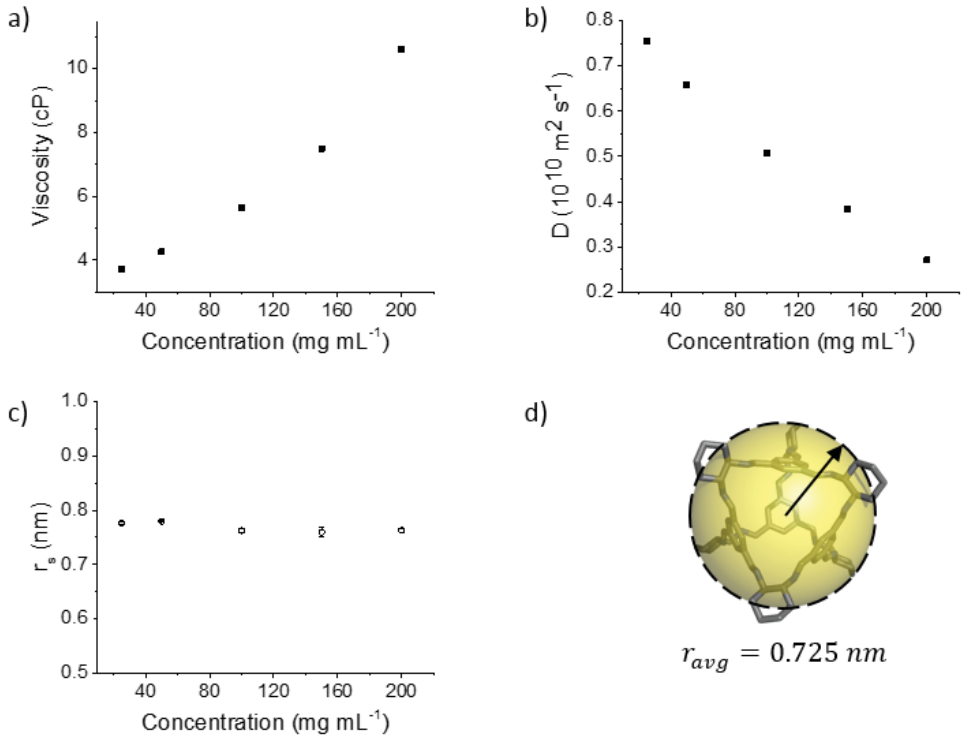


Figure 3.51. (a) Viscosities of PL2 at varying concentrations; (b) Diffusion coefficients obtained by ^{1}H PFG-NMR spectroscopic analysis; (c) Solvodynamic radius, R_H, of cage molecules in solution; (d) The X-Ray crystal structure of CC3. Hydrogens have been omitted for clarity. Sphere represents the average solvodynamic radius measured in (c).

It was noted that r_s remained constant at elevated concentration, and was consistent with a single macromolecular cage (Figure 3.51). Thus, ^{1}H PFG-NMR spectroscopic analysis confirmed that no appreciable aggregation occurred within the time-scale of ^{1}H NMR spectroscopy (here, ~ 10^{-5} s). This was taken to confirm that the chemical shift change associated with dissolution of CH_4 in **PL2** was a result of binding to a single cage cavity, rather than to a cage aggregate.

To assess the strength of cavity binding in **PL2**, it was necessary to perform a number of experiments to correct for both viscosity and proportion of solution occupied by the organic covalent cage (Equation 5, page 6). Estimation of the proportion of time guests are bound, ρ_b, from diffusion NMR has been performed previously.[62] In this case, it was possible to determine the diffusion coefficients of the unbound guest, host and bound guest in a system with constant viscosity. Subsequently, the differences between bound and unbound species were compared to estimate ρ_b (Equation 12).

$$f_b = \frac{D_{guest-observed} - D_{guest-unbound}}{D_{host} - D_{guest-unbound}} \qquad (12)$$

where $D_{guest\text{-}unbound}$ represents the diffusion coefficient of the unbound guest, $D_{guest\text{-}observed}$ represents the diffusion coefficient of the guest when bound, and D_{host} represents the diffusion coefficient of the host molecule, which does not change on binding.

However, because the viscosity of the solution changes during the preparation of **PL2**, the diffusion coefficient of CH_4 in the absence of cage, $D_{guest\text{-}unbound}$, cannot be directly compared to D_{host} and $D_{guest\text{-}observed}$, which are measured in the presence of cage species. It is necessary to correct for this by multiplying each diffusion coefficient by the viscosity of the solution in which it was measured (Equation 13).

$$\rho_b = \frac{(D_{guest-observed} \times \eta_{PL}) - (D_{guest-unbound} \times \eta_{PCP})}{(D_{host} \times \eta_{PL}) - (D_{guest-unbound} \times \eta_{PCP})} \qquad (13)$$

where subscripted η values represent the viscosity of the solution in which the corresponding D were measured. In the case described above where the viscosities of each solution are identical, the added terms cancel, producing Equation 12.

In addition to this, because the concentration of cage in solution is so high, significant increases in the average collision cross-sectional area, σ_{avg} (see Section 2.1.2), were expected, which may cause an apparent decrease in diffusion coefficients

D_{host} and $D_{guest\text{-}observed}$. This again makes any comparison with $D_{guest\text{-}unbound}$, which is measured in neat HCP, less valid. Because the viscosity of **PL2** changes with concentration, it is difficult to apply Equation 5, which would allow this to be corrected.

Instead, measuring $D_{guest\text{-}unbound}$ in the presence of the SIC at an identical mass concentration to **PL2** would replicate the effects of increased σ_{avg} without inducing binding. To confirm that an increase in diffusion coefficient was observed in the presence of the SIC, the relative sizes of several molecules (doped at 10 μL per mL of solution) were measured at an SIC concentration of 200 mg mL^{-1} (Figure 3.52).

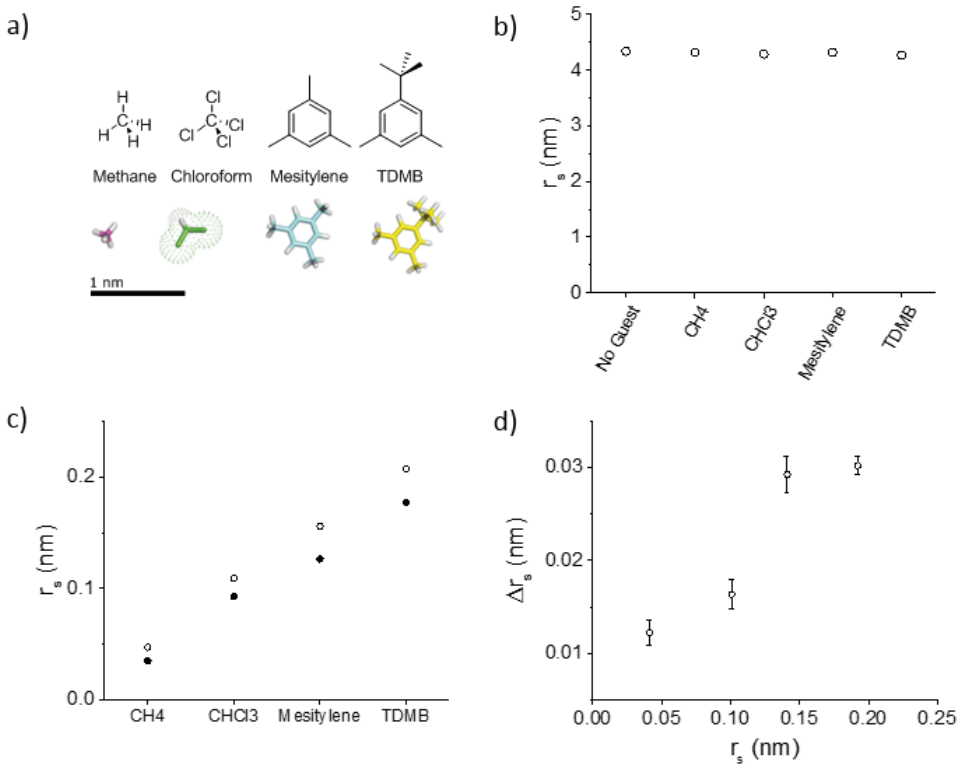


Figure 3.52. ^{1}H PFG-NMR of four additives (10 μL) were measured in neat HCP, and SIC (200 mg mL^{-1}) to estimate the effect of increasing σ_{avg}; (a) Chemical structures of guests alongside molecular models. Scale bar represents 1 nm; (b) Solvodynamic radius of the SIC was measured in each case to confirm viscosities did not change on addition of guest; (c) Solvodynamic radii of guests in neat HCP (empty circles) and the SIC solution (filled circles). Error bars represent 1σ across three measurements.

Changes in solvodynamic radii were found to be small (Figure 3.52c), and correlated with guest size (Figure 3.52d), indicating changes to σ_{avg} were brought about by the presence of the SIC, and not by changes in molecular size.

Finally, to confirm that this represented an adequate approximation of the changes to σ_{avg}, the diffusion coefficient of 1-(*tert*-butyl)-3,5-dimethylbenzene (TDMB) was measured in **PL2.** TDMB had been found to be unable to enter the cage cavity in the original solvent screen.[44] Thus, the D_{obs} should remain unaffected by the presence of

PL2, as no binding would take place. It was thought that this could be used to compare changes to σ_{avg} in **PL2**, compared with the SIC-solution.

To confirm that TDMB was unable to displace CH_4 from **PL2**, Becky Greenaway conducted displacement testing in which TDMB was added to a solution of CH_4-saturated **PL2**, which was observed to release no gas. Subsequent ^{1}H PFG NMR spectroscopic analysis of this mixture demonstrated no chemical shift change in the peak associated with CH_4 (Figure 3.53).

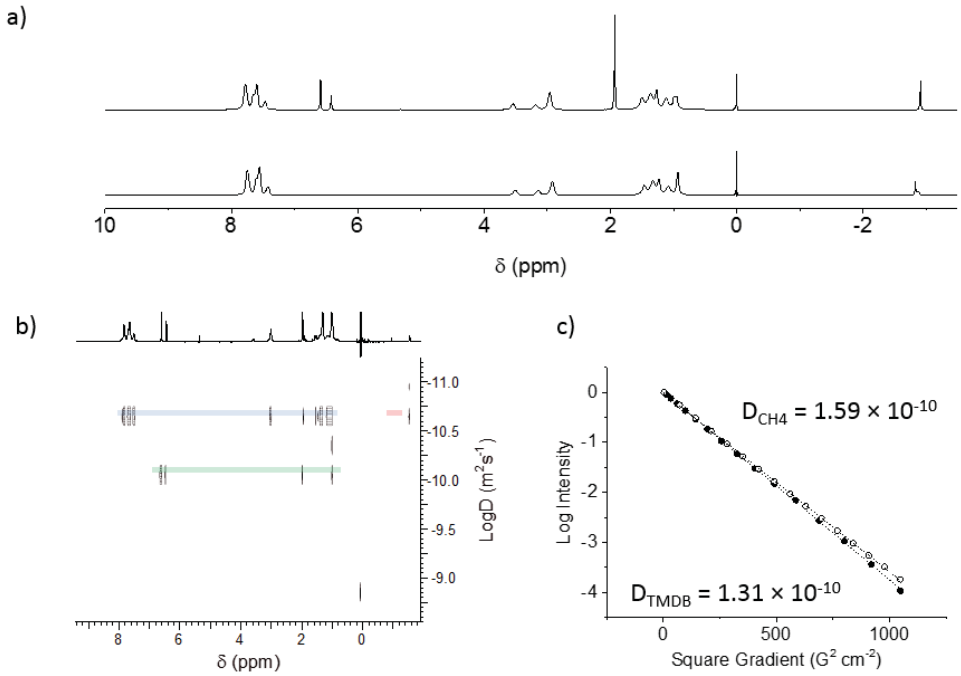


Figure 3.53. (a) ^{1}H NMR spectra of PL2 dosed with CH_4 (top) and once TBMB was added (bottom); (b) ^{1}H DOSY spectrum of PL2 dosed with CH_4 and TBMB. Highlighted: cage (blue), TBMB (green), CH_4 (red); (c) ST-decays for cage and TBMB used to calculate diffusion coefficients.

The diffusion coefficient of CH_4 was observed to decrease from 1.59 to 1.31 × 10^{-11} m^2 cm^{-2}, corresponding to an increase in apparent size. This may be a product of

changing solvent-conditions, which result in an increase in the proportional time CH_4 remains bound, but could also be a result of TDMB inhabiting additional volume in the system (Equation 5).

The solvodynamic radius of TDMB in **PL2** was found to be 1.46 Å, similar to the value of 1.75 Å observed in the SIC-solution. This confirmed that measurements in the SIC-solution were a good approximation for changes in σ_{avg}. On this basis, the value of $D_{guest\text{-}unbound}$ measured in the SIC-solution was used for calculations of proportional occupancy (Figure 3.54)

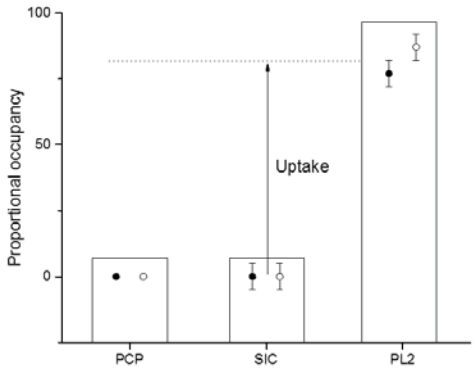


Figure 3.54. Proportional occupancy of guests in PL2, measured by comparing the diffusion coefficient of bound and unbound guests. Methane (filled circles) and chloroform (open circles) are shown. Error bars associated with measurements in PCP are smaller than the size of the data point.

To further assess the validity of these results, selective binding-displacement studies were carried out. In each experiment, CH_4 was dosed into **PL2** and a second guest was added subsequently. Displacement tests by Becky Greenaway confirmed that addition of chloroform to samples of CH_4-saturated **PL2** displaced ~ 80 % of the gas that had been stored in the sample. This was thought to be a result of chloroform binding more strongly to the internal cavity of the cage, which was supported by the

higher proportional occupancy of $CHCl_3$ compared to CH_4. ^1H NMR spectroscopic analysis demonstrated a CH_4 downfield chemical shift change from -2.83 to -0.82 ppm, which indicated that displacement has occurred. This was supported by an 84 % decrease in the integral area of the ^1H NMR peak associated with CH_4, indicating that CH_4 had been displaced from the system.

^1H PFG-NMR spectroscopic analysis of the sample confirmed that r_s of chloroform had increased from 0.93 Å to 4.05 Å, suggesting that it had been incorporated into a larger structure. This was consistent with uptake into the pore of a cage. Concurrently, r_s of CH_4 molecules in solution had decreased from 1.30 Å in the absence of chloroform, to 0.33 Å. This size agreed well with CH_4 in solutions of SIC in hexachloropropene (0.35 Å), which indicated it was probable that the CH_4 diffused through solution, and was no longer bound within the cage (Figure 3.55). In this way, uptake was confirmed and probed by PFG-NMR.

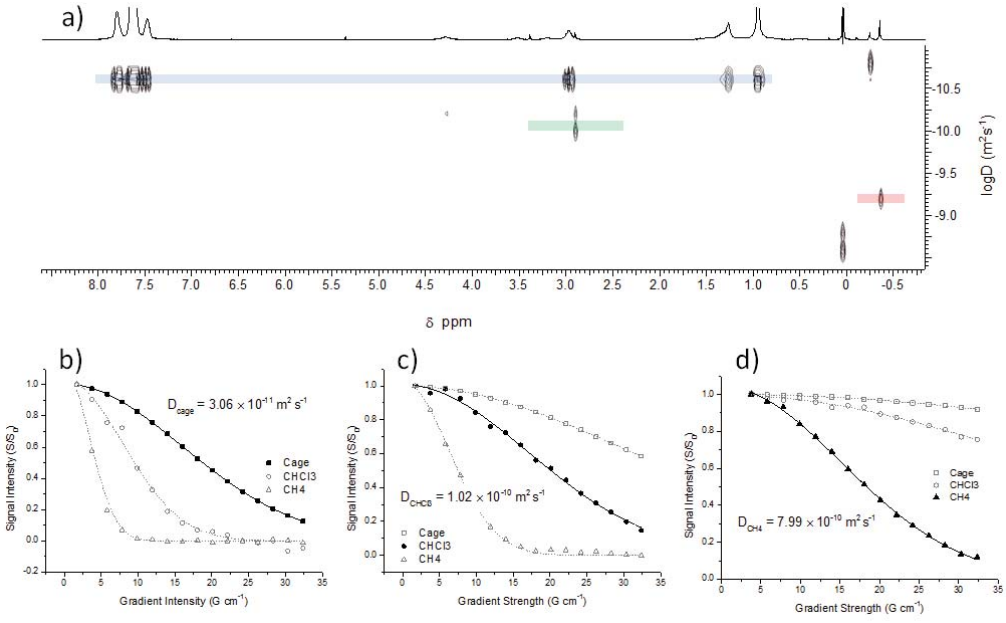


Figure 3.55. (a) ^1H diffusion-ordered (DOSY) spectrum of a PL2 solution containing both CH_4 and $CHCl_3$. Least attenuated ^1H NMR spectrum shown (top); ST-decays for all components are shown for each of three individual experiments that were optimised for (b) cage molecules, (c) chloroform and (d) CH_4.

3.3 Conclusions

In conclusion, four novel dodecaamide cage materials have been synthesised from reactions between acyl chlorides **AC2–4** and **RCC1**. These were designed to measure the effects of acyl chloride bulk on reactivity, but also present the opportunity to extend functionality beyond the cage.

A fifth amide cage, which was synthesised from the bulky **AC1** and **RCC1** was found to substitute only 11 times before becoming kinetically unable to react further. This could be used to generate a cage-monomer, were it possible to functionalise the twelfth amine moiety. However, as access to the nitrogen atoms is limited, this molecule is unlikely to have a permanent pore.

Several attempts were made to desymmetrised **RCC3** and **R1^{1}3^{5}** were made by reaction with acid anhydrides **AA1–4**. In the case of **R1^{1}3^{5}**, significantly enriched samples could be synthesised by addition of stoichiometric quantities of **AA3**.

In the future, it may be possible to combine **R1^{1}3^{5}** with bulkier anhydrides, such as one derived from **AC1**, to further inhibit reactivity rates between anhydrides and reduced, desymmetrised cages. This would offer the opportunity to develop cage polymers, which have programmable distributions of cage material, but retain solution processability.

Finally, PFG-NMR was used to characterise host-guest behaviour in a porous liquid. This technique is well-suited to estimating proportional uptake of guest molecules and could be used to enable further development of porous liquids, as well as investigating guest uptake within cage cavities.

3.4 Experimental

A sealed capillary tube containing $CDCl_3$ and TMS was inserted into NMR tubes along with the sample to facilitate kinetics measurements to be made. As TMS was not present in the $CDCl_3$ used for reactions, the quantity of TMS remained constant, both during and between experiments. In the case of porous liquid PFG-NMR, a sealed lock tube containing 10 % v/v TMS in CD_2Cl_2 was used to facilitate a deuterium lock without affecting the chemical makeup of the porous liquid. All measurements were carried out three times and numbers quoted represent the mean.

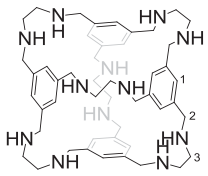
The absolute quantity of TMS seen by the NMR detection coil was calculated by placing the NMR capillary within an NMR tube containing a known concentration of isobutyric anhydride dissolved in $CDCl_3$. By comparing integrations relating to the chemical shifts associated with isobutyric anhydride, with those of the reference TMS, the apparent quantity of TMS was calculated from each capillary.

In subsequent experiments, the peak integrals of reaction products were compared with this known quantity to determine the concentration of product, [*x*], in solution (Equation 14).

$$\frac{\int(x)}{\int(TMS)} \cdot \frac{n_{TMS}}{n_x} \cdot [TMS] = [x] \qquad (14)$$

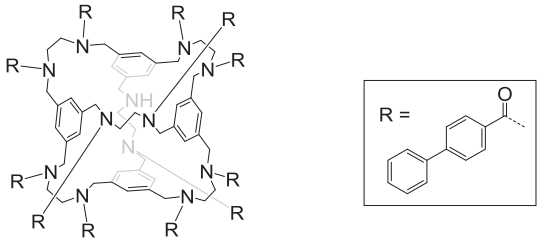
where integrals represent the peak integral associated with a particular chemical shift, *n* indicates the number of protons associated with a particular chemical shift and subscripts indicate the chemical shift under scrutiny. [TMS] represents the known apparent concentration of TMS seen by the NMR detection coil. [*x*] represents the absolute concentration of *x* in solution.

Synthesis of RCC1:



To a stirring solution of **CC1** (500 mg, 0.631 mmol, 1.0 eq.) in MeOH (150 mL) was added portion-wise $NaBH_4$ (770 mg, 20.3 mmol, 32 eq.) and the resultant clear solution was maintained at r.t. for 16 hours. To the solution was added water (2 mL) and the solution was stirred for a further 2 hours. Solvents were removed *in vacuo* to afford an off-white solid. Crude solid was extracted with $CHCl_3$ (200 mL), and the remaining solids were removed by filtration. Organic solvents were removed *in vacuo* to afford a white solid (250 mg, 0.306 mmol, 48%). **MS (ESI)** *m/z* = 817.6 $[M+H]^+$; **^{1}H NMR** ($CDCl_3$, 400MHz) δ ppm 7.07 (12H, s, H1), 3.69 (24H, s, H2), 2.65 (24H, s, H3). Data in accordance with literature values.[13]

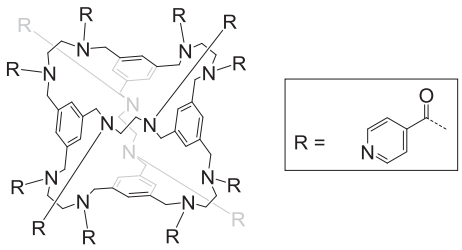
Reaction of RCC1 with [1,1'-biphenyl]-4-carbonyl chloride (RC1j):



To a cooled (0 °C) solution of **RCC1** (82 mg, 0.100 mmol, 1.0 eq.) in $CHCl_3$ (10 mL) was added a solution of [1,1'-biphenyl]-4-carbonyl chloride (303 mg, 1.40 mmol, 14.0 eq.) and NEt_3 (204 μL, 1.50 mmol, 15.0 eq.) to afford a clear colourless solution, which was stirred for 16 hours. To the solution, after cooling (0 °C), was added *aq*. NaOH (1 M, 20 mL). The resultant biphasic mixture was separated and the organic solvents were washed with brine (50% sat., 2 × 10 mL). After the aqueous layer was extracted with $CHCl_3$, organic solvents were combined, dried over $MgSO_4$,

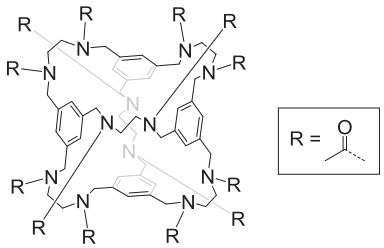
and removed *in vacuo* to afford a white powder (155 mg, 0.055 mmol, 55 %). **IR** 3449.5, 2921.0, 2347.1, 1629.4 cm^{-1}; **HR-MS (ESI)** $C_{191}H_{162}N_{12}O_{11}^{2+}$ calc. 1400.6271 expt. 1400.6256, **^{1}H NMR** (chloroform-d, 400 MHz) δ ppm 7.74–7.31 (m, Ar*H*, 109H), 7.25–6.88 (m, Ar*H*, 12H) 5.17–4.58 (m, ArC*H*$_2$N, 24H), 4.19–3.10 (m, C*H*$_2$, 24H); **^{13}C NMR** (101 MHz) δ 172.9, 142.6, 140.1, 138.4, 134.6, 127.8, 52.0, 38.6.

Reaction of RCC1 with isonicotinoyl chloride (RC1k)



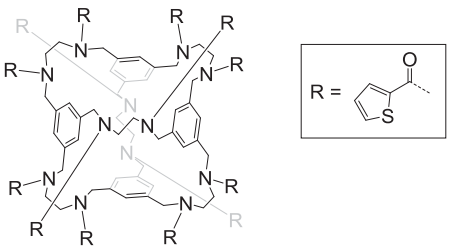
To a cooled (0 °C) solution of **RCC1** (82 mg, 0.100 mmol, 1.0 eq.) in $CHCl_3$ (10 mL) was added a solution of isonicotinoyl chloride hydrochloride (249 mg, 1.40 mmol, 14.0 eq.) and NEt_3 (408 μL, 3.00 mmol, 30.0 eq.) to afford a pale yellow solution, which was stirred for 16 hours. To the solution, after cooling (0 °C), was added *aq*. NaOH (1 M, 20 mL). The resultant biphasic mixture was separated and the organic solvents were washed with brine (50% sat., 2 × 10 mL). After the aqueous layer was extracted with $CHCl_3$, organic solvents were combined, dried over $MgSO_4$, and removed *in vacuo* to afford a light yellow solid (146 mg, 0.070 mmol, 70 %). **IR** 2926.5, 2345.6, 1629.2 cm^{-1}; **MS (ESI)** m/z = 2078.9 $[M+H]^+$; **^{1}H NMR** (chloroform-d, 400 MHz) δ ppm 8.63 (24H, d, *J* = 5.0 Hz, H5), 7.83 (24H, d, *J* = 5.0 Hz, H2), 6.99 (12H, s, H1), 4.68 (24H, br s, H2), 3.59 (24H, br s, H3); **^{13}C NMR** (101 MHz) δ 151.3, 147.6, 134.5, 131.4, 123.8, 93.8, 87.6, 46.3.

Reaction of RCC1 with acetyl chloride (RC1l)



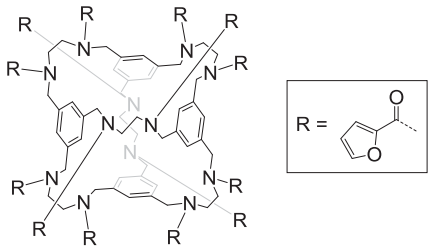
An identical method was used as that employed in the synthesis of **RCC1b**, using 4-bromobutanoyl chloride (260 mg, 1.50 mmol, 15.0 eq.) to afford a light orange solid (201 mg, 0.072 mmol, 71 %). **IR** 3430.5, 2934.0, 1622.3 cm^{-1}; **HR-MS (ESI)** $C_{72}H_{97}N_{12}O_{12}^{+}$ calc. 1321.7343 expt. 1321.7658, **^{1}H NMR** (chloroform-d, 400 MHz) δ ppm 7.23–6.37 (br. s, 12H, Ar*H*), 5.54–4.00 (br. s, 24H, NC*H*$_2$Ar), 4.00–3.12 (br. s, 24H, NC*H*$_2$) 2.42–1.89 (m, -C(O)C*H*$_3$, 36H); **^{13}C NMR** (101 MHz) δ 171.5, 138.0, 125.0, 53.9, 47.8, 45.8, 30.9, 29.7, 21.4, 9.0.

Reaction of RCC1 with thiophene-2-carbonyl chloride (RC1m)



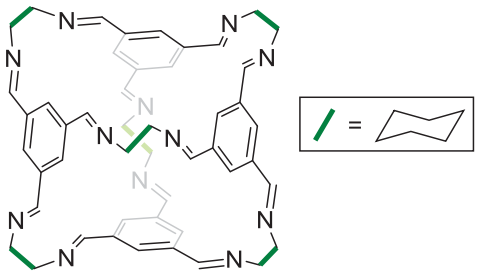
An identical method was used as that employed in the synthesis of **RCC1b**, using thiophene-2-carbonyl chloride (150 μL, 1.50 mmol, 15.0 eq.) to afford an off white solid (165 mg, 0.0587 mmol, 59 %). **IR** 3092.2, 2931.6, 1738.4, 1609.5; **HR-MS (ESI)** $C_{108}H_{96}N_{12}Na_2O_{12}S_{12}^{2+}$ calc. 1091.6869 expt. 1092.6879, **^{1}H NMR** (chloroform-d, 400 MHz) δ ppm 7.75–6.55 (br. s, 24H, Ar*H*), 6.55–6.03 (br. s, 24H, Ar*H*), 5.54–4.26 (br. s, NC*H*$_2$Ar, 24H), 3.05–4.25 (br. s, C*H*$_2$, 24H); **^{13}C NMR** (101 MHz) δ 172.5, 151.8, 144.1, 111.5, 48.2, 27.5.

Reaction of RCC1 with furan-2-carbonyl chloride (RC1n)



An identical method was used as that employed in the synthesis of **RCC1b**, using furan-2-carbonyl chloride (183 mg, 1.50 mmol, 15.0 eq.) to afford a light brown solid (211 mg, 0.075 mmol, 75 %). **IR** 3135.0, 2934.1, 1601.8 cm^{-1}; **HR-MS (ESI)** $C_{108}H_{96}N_{12}Na_2O_{24}^{2+}$ calc. 995.8239 expt. 995.8252, **^{1}H NMR** (chloroform-d, 400 MHz) δ ppm 7.74–7.13 (m, Ar*H*, 36H), 7.12–6.75 (br. s, Ar*H*, 12H), 5.25–4.50 (br. s, NCH_2Ar, 24H), 3.25–4.13 (br. s, CH_2, 24H); **^{13}C NMR** (101 MHz) δ 171.5, 138.0, 125.0, 53.9, 47.7, 45.8, 30.9, 29.7, 21.4, 9.0.

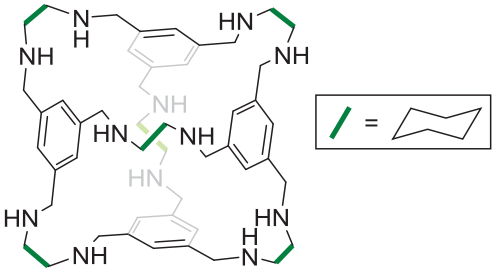
Synthesis of CC3:



Over solid benzene-1,3,5-tricarboxaldehyde (5.00 g, 30.8 mmol, 4 eq.) was layered DCM (100 mL), and subsequently a solution of (R,R)-cyclohexane-1,2-diamine (5.28 g, 46.3 mmol, 6 eq.) in DCM (120 mL). To the resultant mixture was added dropwise trifluoroacetic acid (5 drops), and the mixture was allowed to stand for 14 days. The resultant dark orange solution was filtered and the filtrate discarded. The remaining light yellow solids were washed with a 95:5 mixture of ethanol and DCM (200 mL). After washing the product was dissolved in a fresh 95:5 mixture of ethanol

and DCM (200 mL) by sonication and solids were removed by filtration. Organic solvents were removed under reduced pressure to afford an off-white solid (6.31 g, 5.65 mmol, 63 %). **MS (ESI)** m/z = 1118.5 $[M+H]^+$; **^{1}H NMR** (chloroform-d, 400 MHz) δ ppm 1.34 - 1.90 (m, C*H*$_2$, 48H) 3.25 - 3.39 (m, C*H*N, 12H) 7.89 (s, Ar*H*, 12 H) 8.16 (s, C*H*=N, 12H). Data in accordance with literature values.[15]

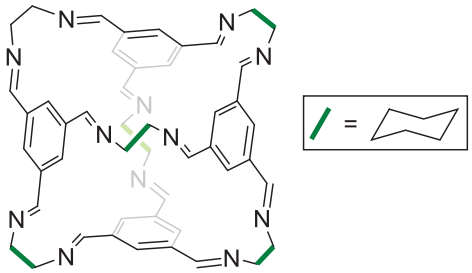
Synthesis of RCC3:



Over solid benzene-1,3,5-tricarboxaldehyde (3.00 g, 18.5 mmol, 4 eq.) was layered DCM (750 mL), and subsequently a solution of (R,R)-cyclohexane-1,2-diamine (3.17 g, 27.8 mmol, 6 eq.) in DCM (120 mL). To the resultant mixture was added dropwise trifluoroacetic acid (15 drops), and the mixture was allowed to stand for 8 days. The resultant orange solution was reduced to dryness under reduced pressure to afford alight orange solid (9.72 g). This was suspended in MeOH (700 mL) and to this suspension was added $NaBH_4$ (4.00 g, 106 mmol, 23 eq.). After 16 h, solvents were removed under reduced pressure to afford an off-white solid, which was suspended in chloroform (150 mL). Solids were removed by filtration (0.7 μm mesh) and discarded. Organic solvents were removed under reduced pressure to afford a light orange foam (7.64 g). Acetone was added and the mixture was stirred until all solids had dissolved. A white, shiny precipitate was formed on standing for 7 days, which was separated by filtration. Solids were dissolved in a 1:1 mixture of MeOH and $CHCl_3$ (400 mL) and water (5 mL) was added. After 24 hours, solvents were removed under reduced

pressure to afford a white solid (3.71 g, 3.25 mmol, 70 %). **MS (ESI)** m/z = 1141.8 $[M+H]^+$, **^{1}H NMR** (500 MHz, chloroform-*d*) δ ppm 0.96 (d, *J*=8.4 Hz, 12 H) 1.09 - 1.22 (m, 12 H) 1.65 (d, *J*=7.8 Hz, 12 H) 2.00 (d, *J*=12.8 Hz, 12 H) 2.15 - 2.23 (m, 12 H) 3.70 (dd, *J*=126.6, 14.0 Hz, 24 H) 7.13 (s, 12 H). Data in accordance with literature values.[13]

Synthesis of 1^13^5:



To a stirring solution of 1,3,5-tricarboxaldehyde (1.16 g, 7.15 mmol, 4eq.) in DCM (500 mL) was added a solution of ethylenediamine-^{15}N dihydrochloride (239.5 mg, 1.77 mmol, 1 eq.) and NEt_3 (0.493 mL, 3.54 mmol, 2 eq.) in DCM (100 mL), and a solution of (1R,2R)-cyclohexane-1,2-diamine (1.01 g, 8.85 mmol, 5 eq.) in DCM (100 mL). The resultant solution was stirred at r.t. for 3 d. Organic solvents were removed under reduced pressure to afford a light orange solid (2.86 g). Crude material was purified by preparative HPLC using a Synchronis C8 column (particle size 5 μm, column size: 150 mm × 46 mm), with a column oven temperature of 30 °C and a flow rate of 1 mL min^{-1}. Fractions were collected between 4.8 and 5.4 minutes (Figure 3.56); pure fractions were combined and organic solvents were removed under reduced pressure to afford a white solid (789 mg, 0.742 mmol, 42 %). **HR-MS (ESI)** $C_{68}H_{103}N_{12}^+$ expt. 1087.8422, calc. 1087.8429. **^{1}H NMR** (500 MHz, chloroform-*d*) δ

ppm 1.40 - 1.55 (12 H, m) 1.58 - 1.77 (23 H, m) 1.83 (13 H, d, *J*=7.93 Hz) 3.26 - 3.42 (11 H, m) 3.79 (2 H, d, *J*=9.31 Hz) 4.16 - 4.28 (2 H, m) 7.83 - 7.95 (14 H, m) 8.12 - 8.19 (14 H, m), **^{13}C NMR** (126 MHz) δ 24.35, 32.98, 61.46, 74.60, 74.80, 129.30, 129.43, 129.54, 129.64, 136.22, 136.27, 136.57, 136.71, 158.91, 158.97, 159.04, 159.14, 159.20, 161.64, 161.69, **^{15}N NMR** (51 MHz) δ 326.30.

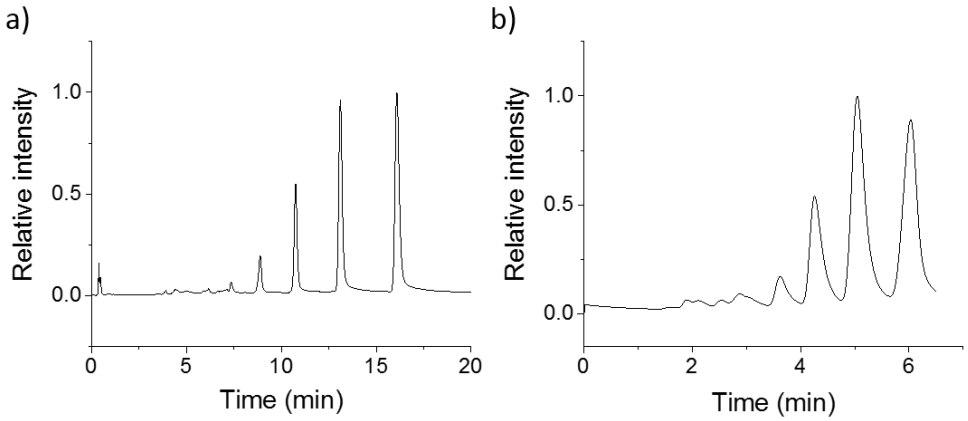
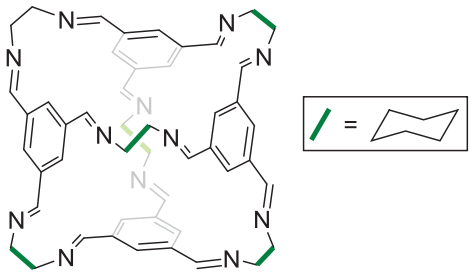


Figure 3.56. HPLC trace of mixtures of desymmetrised cages using (a) analytical and (b) preparative Synchronis C8 columns.

Synthesis of {15N}R1^{1}3^{5}:



To a stirring solution of 1^{1}3^{5} (398 mg, 0.38 mmol, 1 eq.) in a 9:1 mixture of methanol and chloroform (150 mL) was added portionwise $NaBH_4$ (600 mg, 15.7 mmol, 42 eq.) and the resultant clear solution was stirred at room temperature for 16 hours. Water

(5 mL) was adde to quench $NaBH_4$ that remained and after 4 hours, solvents were removed under reduced pressure to afford a white solid, which was suspended in chloroform (150 mL). Solids were removed by filtration (0.7 μm mesh) and discarded. Organic solvents were removed under reduced pressure to afford a white foam (361 mg, 0.30 mmol, 80 %). **HR-MS (ESI)** $C_{68}H_{103}N_{12}^{+}$ expt. 1087.8422, calc. 1087.8429. **CHN** calculated for $C_{68}H_{103}N_{12}Na$: C 73.54 %, H 9.26 %, N 14.03. Found: C 73.54 %, H 9.12 %, N 14.03 %. **^{1}H NMR** (500 MHz, chloroform-*d*) δ ppm 7.30 (s, 2H, Ar*H*) 7.27 (s, 2H, Ar*H*), 7.12 (s, 2H, Ar*H*), 7.08 (s, 2H, Ar*H*) , 6.90 (s, 2H, Ar*H*), 6.87 (s, 2H, Ar*H*), 4.13–3.74 (m, NCH_2Ar, 10H), 3.66–3.28 (m, NCH_2Ar, 14H), 2.30 – 1.94 (m, 14H, C*H* on cyclohexane and CH_2 ethylenediamine), 1.32 – 0.60 (m, 40H, CH_2 on cyclohexane). **^{13}C NMR** (126 MHz) δ 141.3, 141.2, 141.0, 140.9, 125.9, 125.6, 125.5, 125.4, 125.2, 125.1, 61.5, 61.4, 61.2, 60.8, 60.7, 53.4, 51.5, 50.6, 50.2, 50.1, 31.9, 31.8, 31.7, 25.1, 25.05, 25.0, 24.9, 24.8. **^{15}N NMR** (51 MHz) δ 37.99.

3.4.1 NMR Spectra of novel compounds.

Spectra for RC1j:

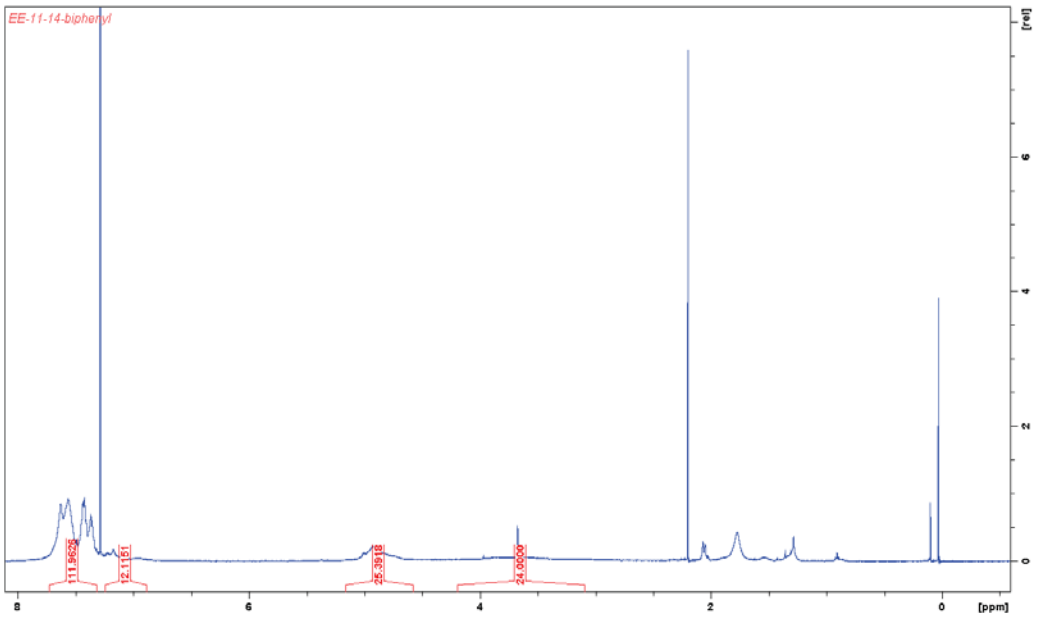


Figure 3.57. ^{1}H NMR Spectrum of RC1j

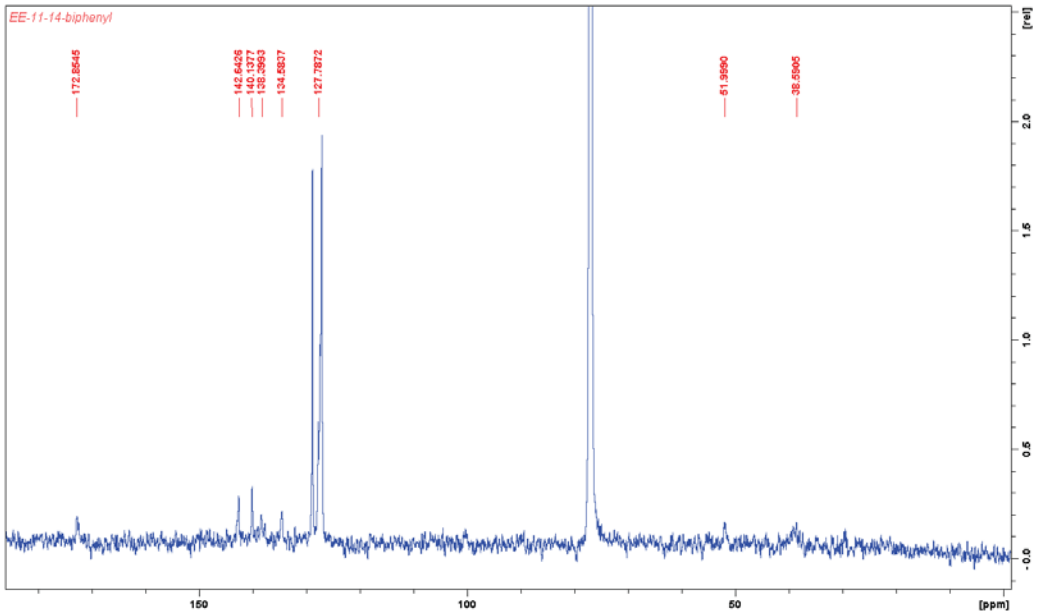


Figure 3.58. ^{13}C NMR Spectrum of RC1j

Spectra for RC1k:

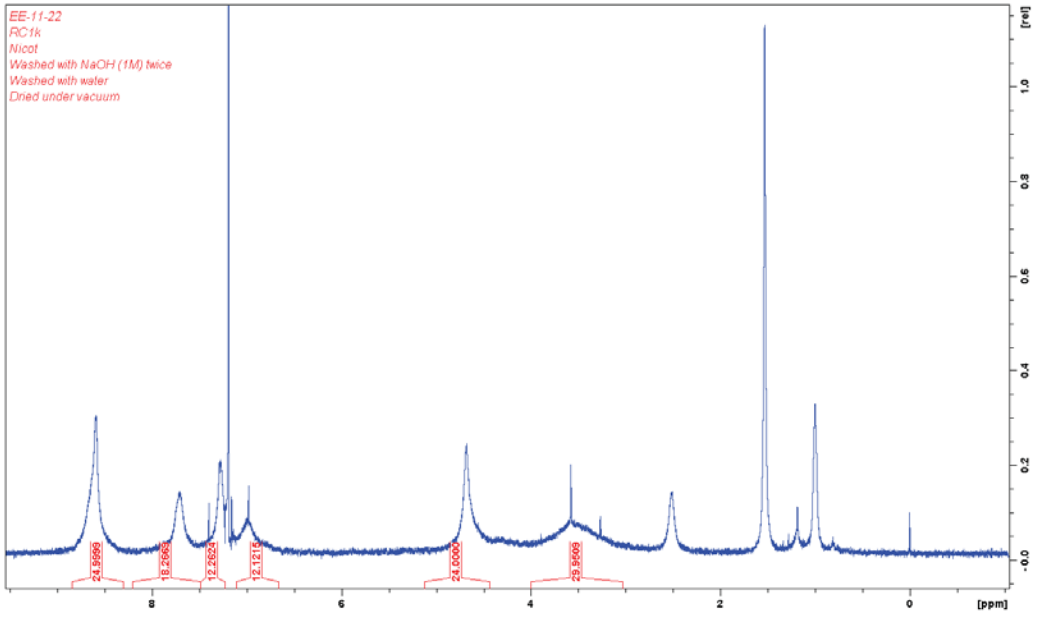


Figure 3.59. ^{1}H NMR Spectrum of RC1k

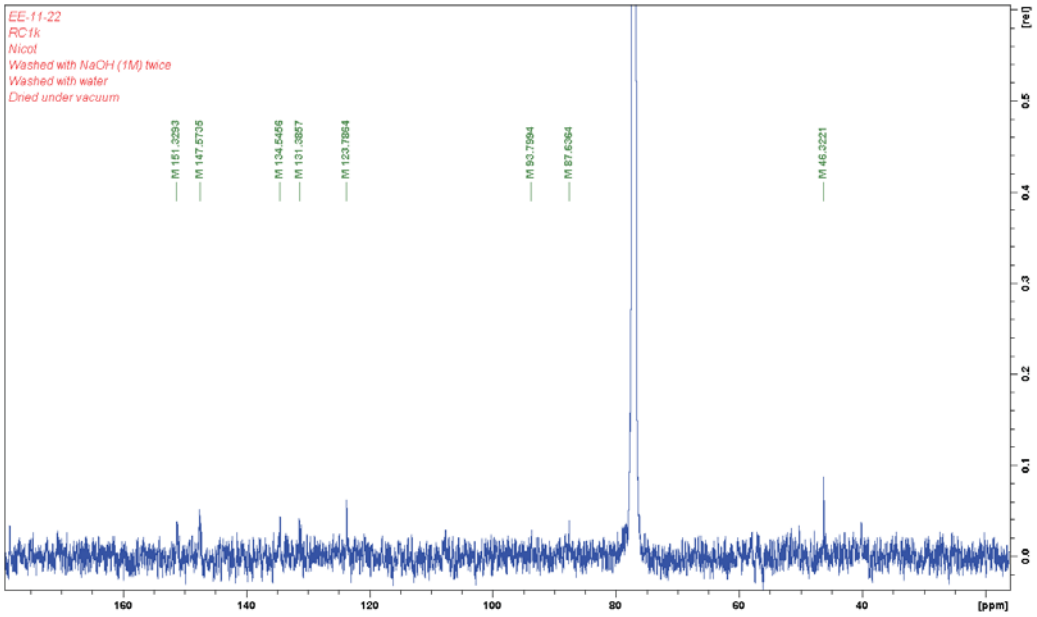


Figure 3.60. ^{13}C NMR Spectrum of RC1k

Spectra for RC1l:

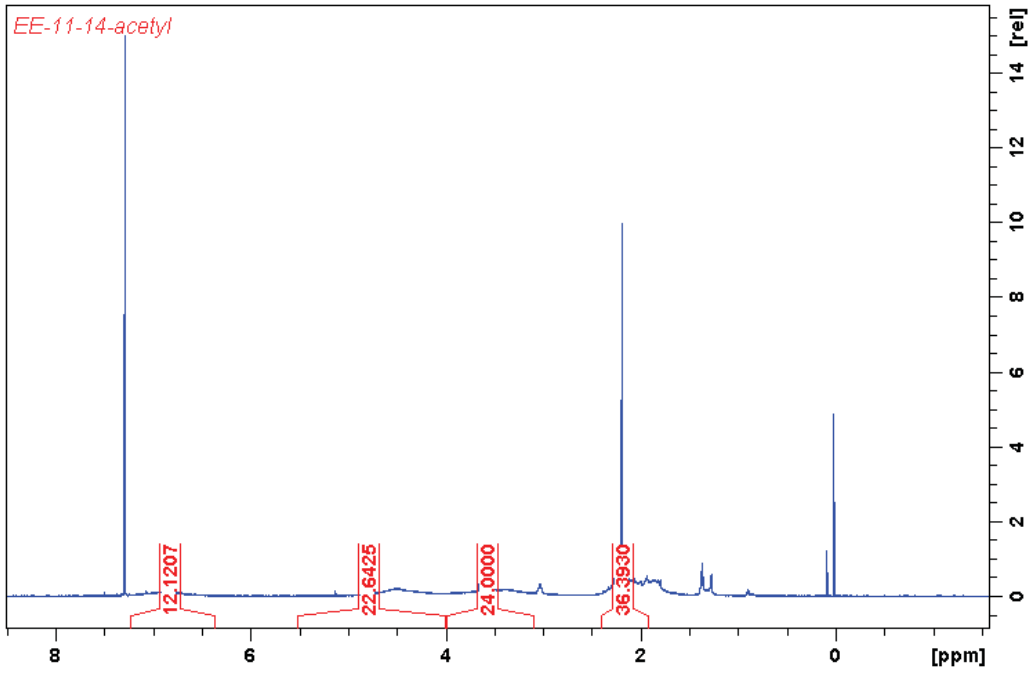


Figure 3.61. ^{1}H NMR Spectrum of RC1l

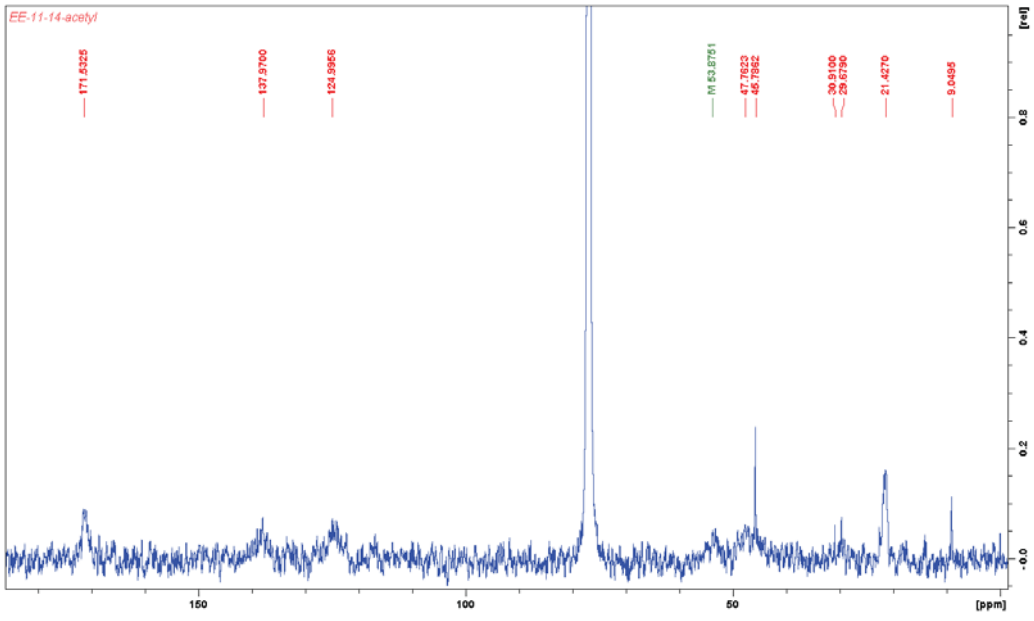


Figure 3.62. ^{13}C NMR Spectrum of RC1l

Spectra for RC1m:

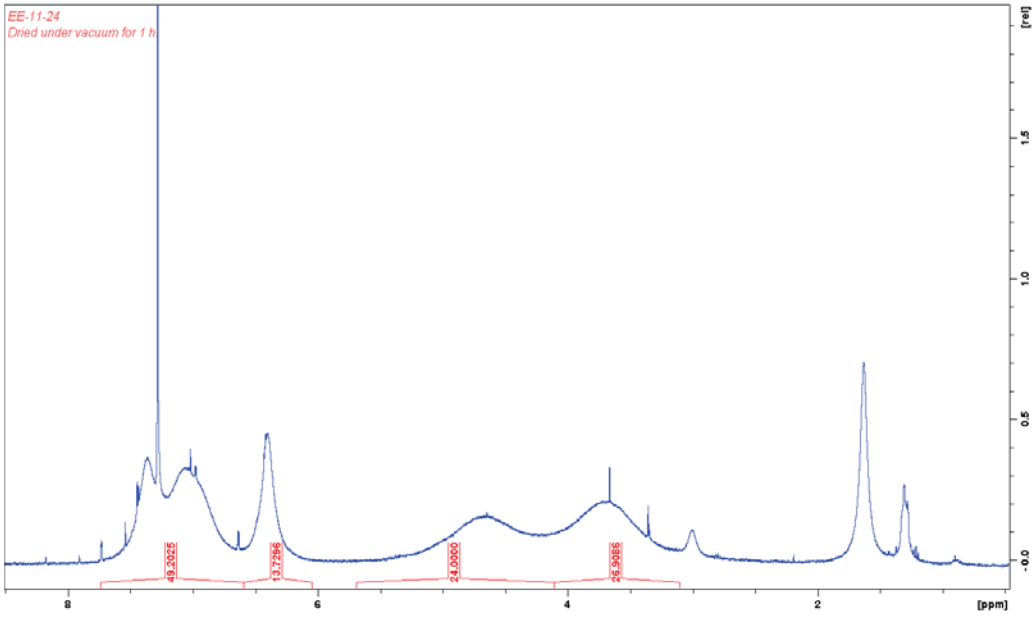


Figure 3.63. ^{1}H NMR Spectrum of RC1m

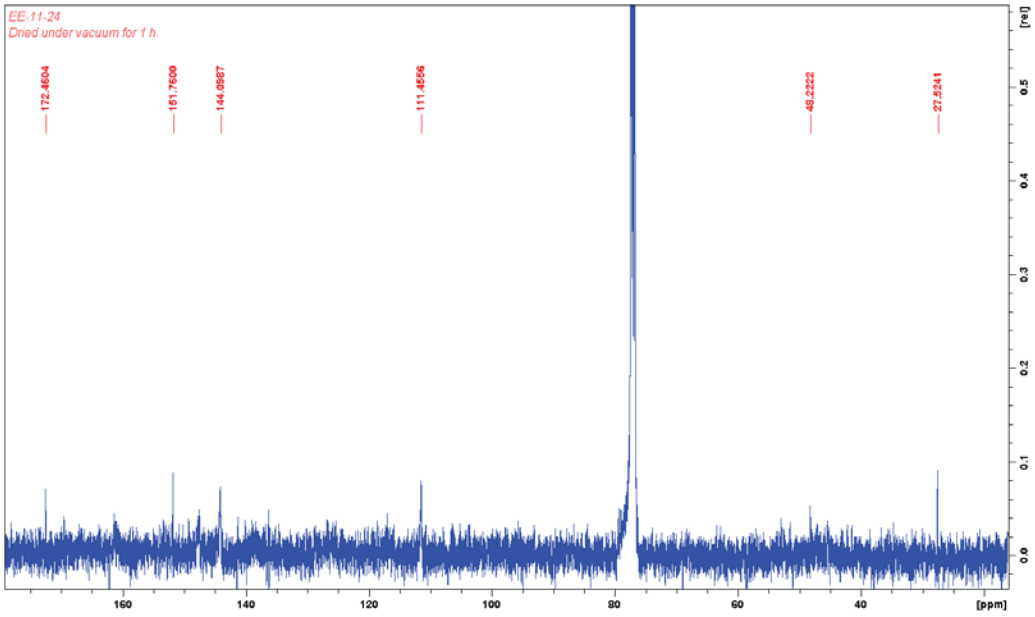


Figure 3.64. ^{13}C NMR Spectrum of RC1m

Spectra for RC1n:

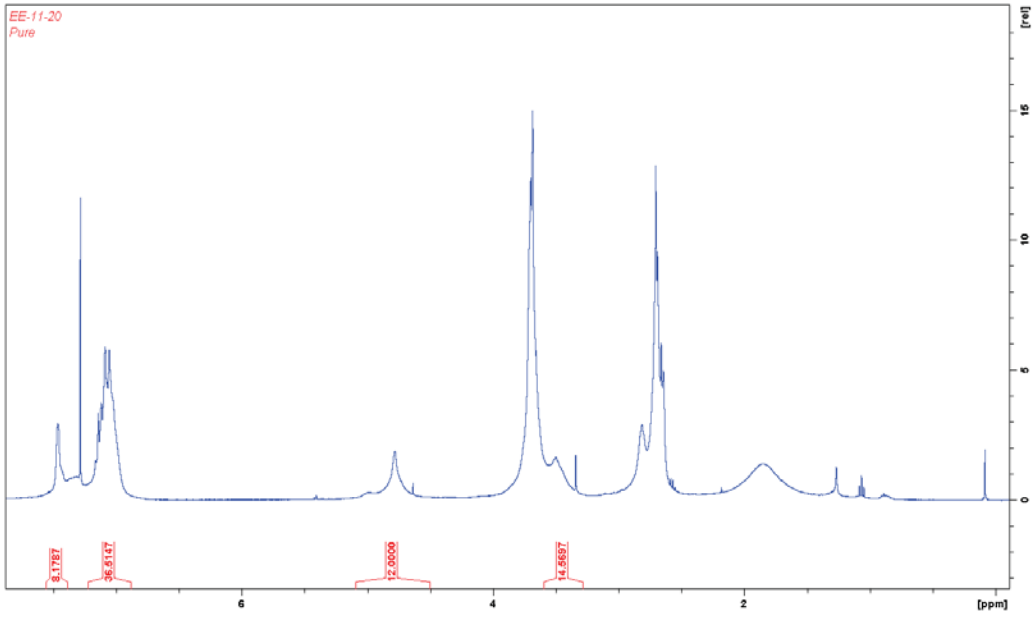


Figure 3.65. ^{1}H NMR Spectrum of RC1n

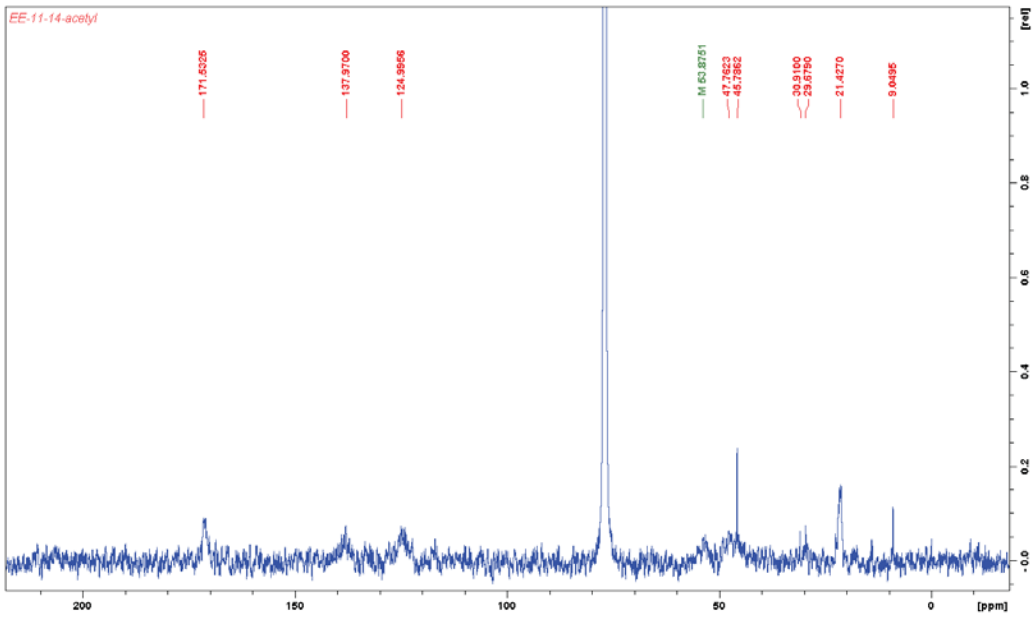


Figure 3.66. ^{13}C NMR Spectrum of RC1n

Spectra for $1^{1}3^{5}$:

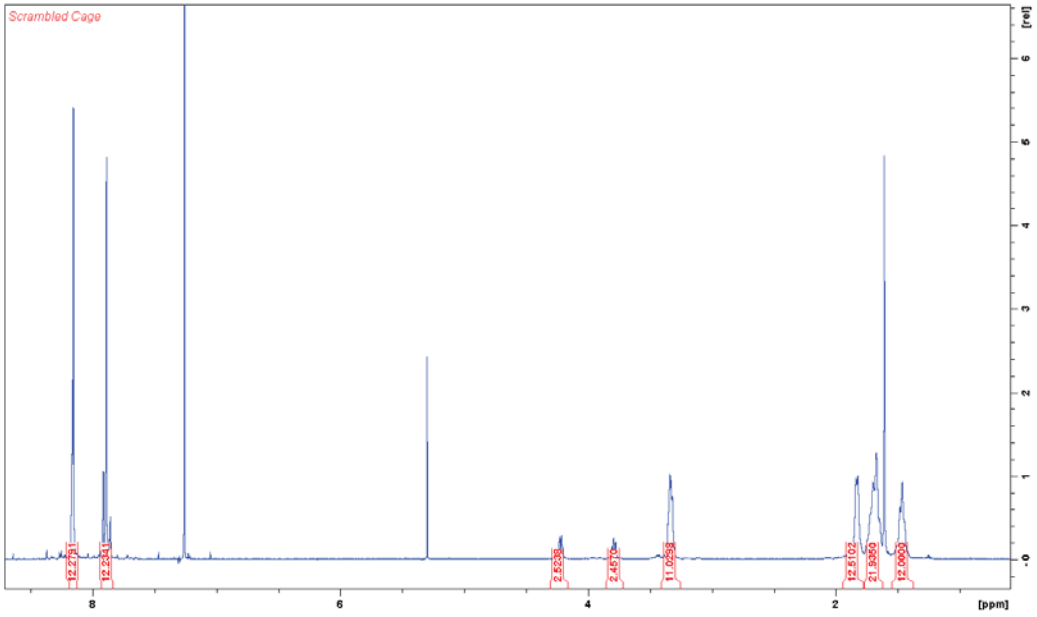


Figure 3.67. ^{1}H NMR Spectrum of $1^{1}3^{5}$

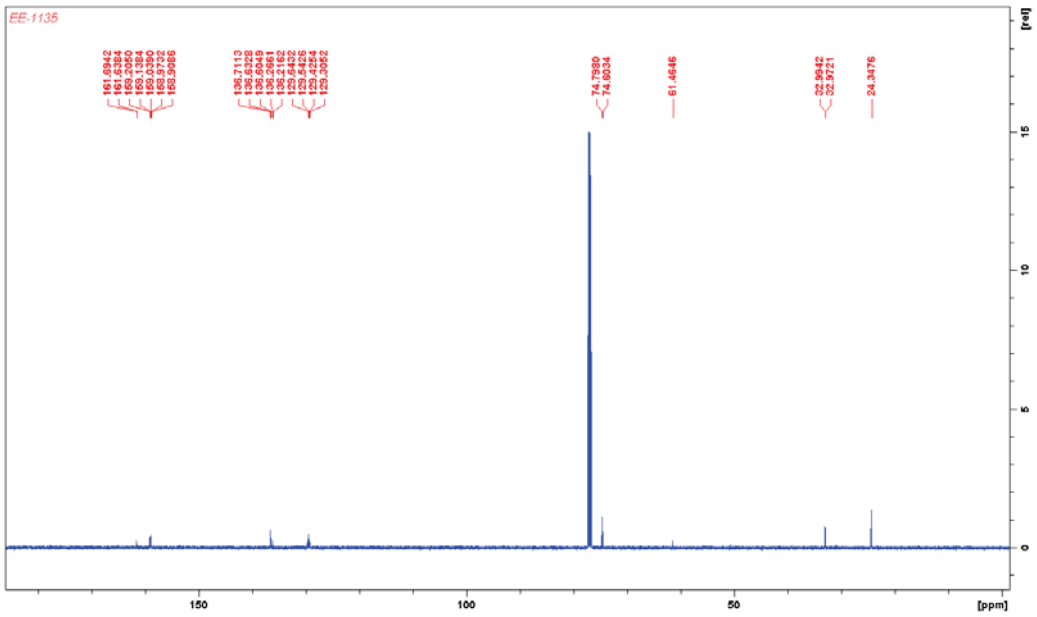


Figure 3.68. ^{13}C NMR Spectrum of $1^{1}3^{5}$

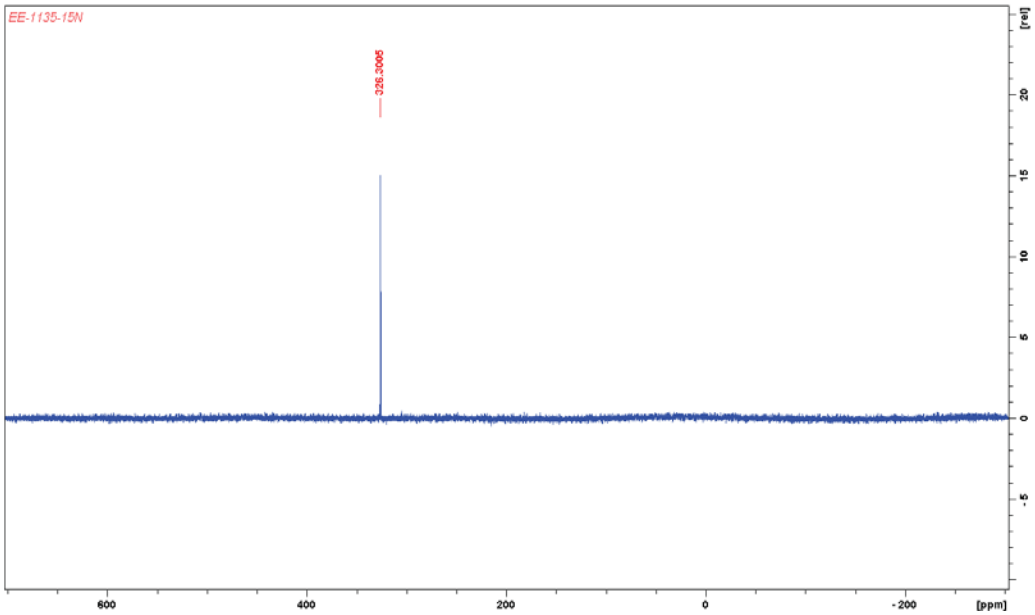


Figure 3.69. ^{15}N NMR Spectrum of {^{15}N}R1^{1}3^{5}

Spectra for R1^{1}3^{5} and {15N}R1^{1}3^{5}:

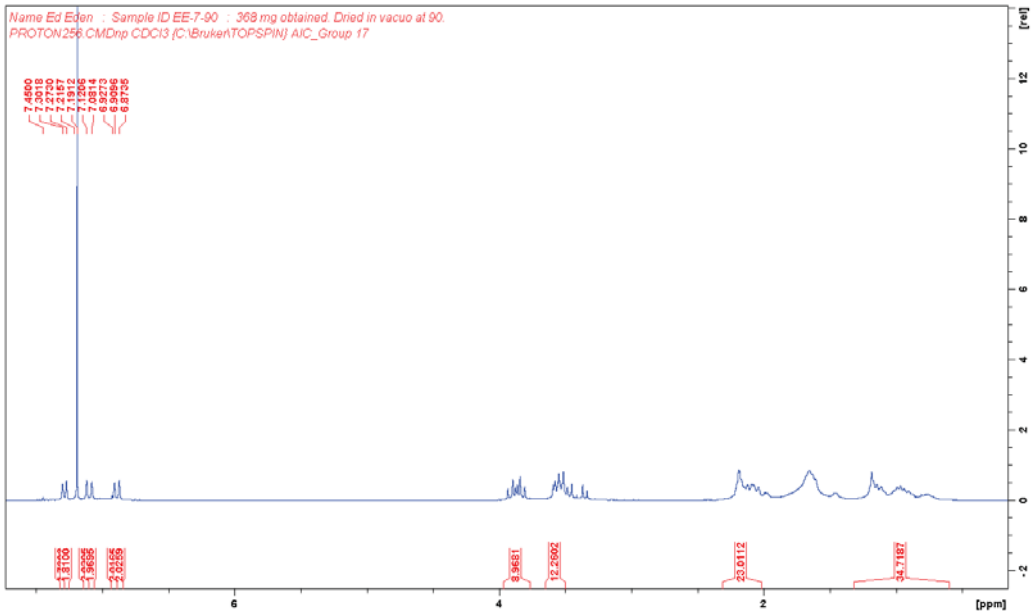


Figure 3.70. ^{1}H NMR Spectrum of R1^{1}3^{5}

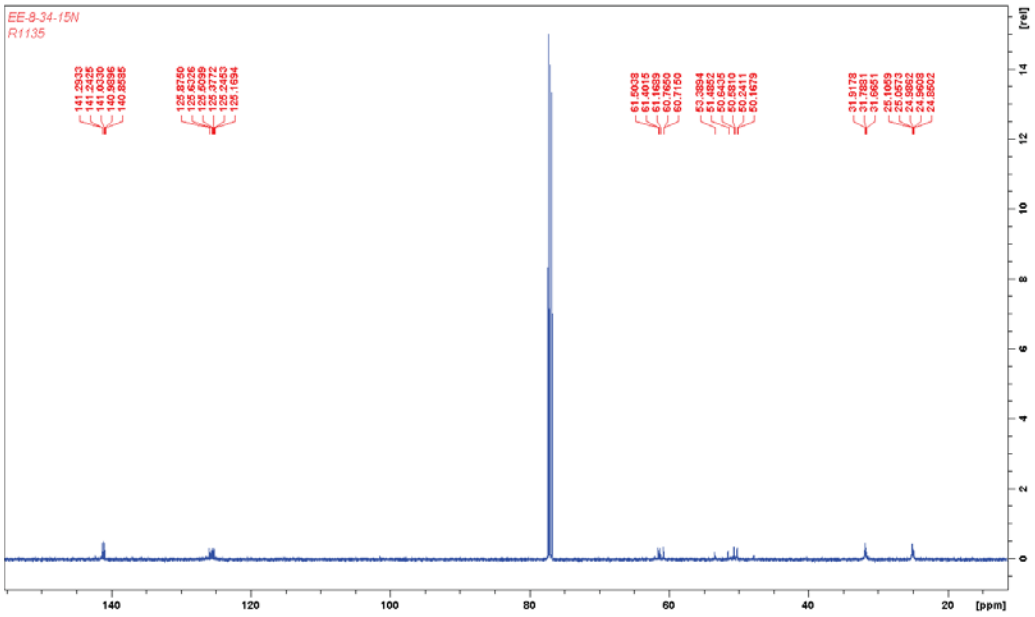


Figure 3.71. ^{13}C NMR Spectrum of R1[1]3[5]

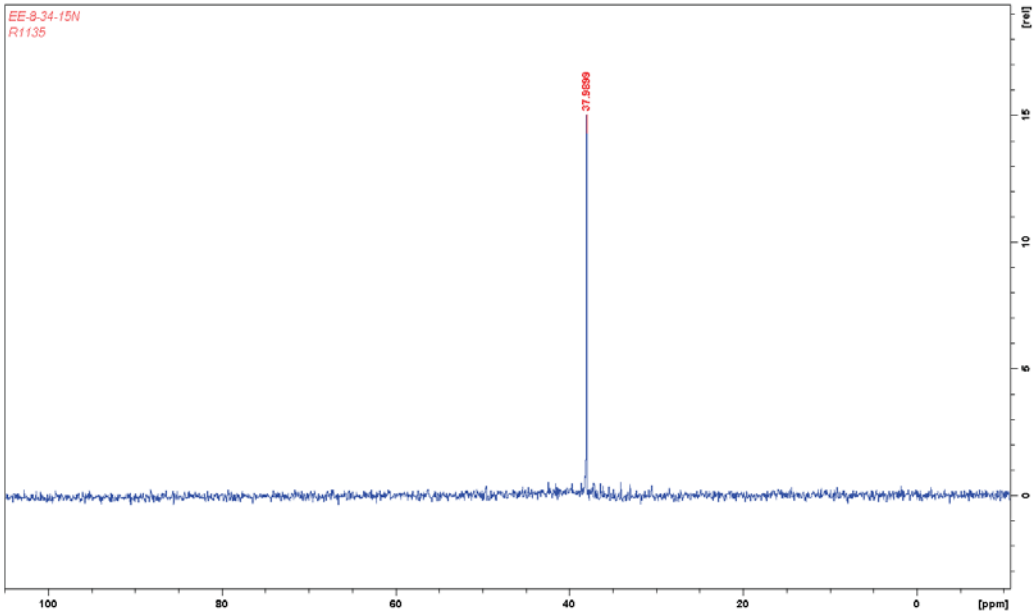


Figure 3.72. ^{15}N NMR Spectrum of {^{15}N}R1[1]3[5]

3.5 Bibliography

1 L. Avram and Y. Cohen, *Chem. Soc. Rev.*, 2015, **44**, 586–602.

2 Q. Song, S. Jiang, T. Hasell, M. Liu, S. Sun, A. K. Cheetham, E. Sivaniah and A. I. Cooper, *Adv. Mater.*, 2016, **28**, 2629–2637.

3 M. Brutschy, M. W. Schneider, M. Mastalerz and S. R. Waldvogel, *Adv. Mater.*, 2012, 6049–6052.

4 M. Drobek, M. Bechelany, C. Vallicari, A. Abou Chaaya, C. Charmette, C. Salvador-Levehang, P. Miele and A. Julbe, *J. Memb. Sci.*, 2015, **475**, 39–46.

5 S. Qiu, M. Xue and G. Zhu, *Chem. Soc. Rev.*, 2014, **43**, 6116–6140.

6 M. Bechelany, M. Drobek, C. Vallicari, A. Abou Chaaya, A. Julbe and P. Miele, *Nanoscale*, 2015, **7**, 5794–5802.

7 D. Umeyama, S. Horike, M. Inukai, T. Itakura and S. Kitagawa, *J. Am. Chem. Soc.*, 2015, **137**, 864–870.

8 R. L. Hartman and H. S. Fogler, *Langmuir*, 2007, **23**, 5477–5484.

9 N. B. McKeown and P. M. Budd, *Macromolecules*, 2010, **43**, 5163–5176.

10 P. M. Budd, B. S. Ghanem, S. Makhseed, N. B. McKeown, K. J. Msayib and C. E. Tattershall, *Chem. Commun.*, 2004, **11**, 230.

11 S. Kitagawa, R. Kitaura and S. Noro, *Angew. Chem. Int. Ed.*, 2004, **43**, 2334–2375.

12 J. R. Holst, A. Trewin and A. I. Cooper, *Nat. Chem.*, 2010, **2**, 915–920.

13 S. I. Swamy, J. Bacsa, J. T. a Jones, K. C. Stylianou, A. Steiner, L. K. Ritchie, T. Hasell, J. a Gould, A. Laybourn, Y. Z. Khimyak, D. J. Adams, M. J. Rosseinsky and A. I. Cooper, *J. Am. Chem. Soc.*, 2010, **132**, 12773–12775.

14 A. F. Bushell, P. M. Budd, M. P. Attfield, J. T. a. Jones, T. Hasell, A. I. Cooper, P. Bernardo, F. Bazzarelli, G. Clarizia and J. C. Jansen, *Angew. Chem. Int. Ed.*, 2012, **52**, 1253–1256.

15 T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S. Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner and A. I. Cooper, *Nat. Mater.*, 2009, **8**, 973–978.

16 K. E. Jelfs, X. Wu, M. Schmidtmann, J. T. A. Jones, J. E. Warren, D. J. Adams and A. I. Cooper, *Angew. Chem. Int. Ed.*, 2011, **50**, 10653–10656.

17 M. Mastalerz, *Synlett*, 2013, **24**, 781–786.

18 J. L. Culshaw, G. Cheng, M. Schmidtmann, T. Hasell, M. Liu, D. J. Adams and A. I. Cooper, *J. Am. Chem. Soc.*, 2013, **135**, 10007–10010.

19 J. Cai, Y. Xing and X. Zhao, *RSC Adv.*, 2012, **2**, 8579.

20 S. Jiang, J. T. A. Jones, T. Hasell, C. E. Blythe, D. J. Adams, A. Trewin and A. I. Cooper, *Nat. Commun.*, 2011, **2**, 207.

21 D. J. Cram and J. M. Cram, *Container Molecules and Their Guests*, Royal Society of Chemistry, Cambridge, 1997.

22 P. Mateus, R. Delgado, P. Brandão and V. Félix, *J. Org. Chem.*, 2009, **74**, 8638–46.

23 S. Kubik, *Angew. Chem. Int. Ed.*, 2009, **48**, 1722–5.

24 P. Mateus, R. Delgado, P. Brandão, S. Carvalho and V. Félix, *Org. Biomol. Chem.*, 2009, **7**, 4661–73.

25 Y. Jin, B. A. Voss, A. Jin, H. Long, R. D. Noble and W. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 6650–6658.

26 G. R. Desiraju, *J. Mol. Struct.*, 2003, **656**, 5–15.

27 F. C. Pigge, *CrystEngComm*, 2011, **13**, 1733.

28 F. G. Bordwell, *Acc. Chem. Res.*, 1988, **21**, 456–463.

29 F. G. Bordwell and H. E. Fried, *J. Org. Chem.*, 1981, **46**, 4327–4331.

30 R. M. Silverstein, *Spectrometric Identification of Organic Compounds*, Wiley, New York, Fourth., 1981.

31 W. A. Henderson and C. J. Schultz, *J. Org. Chem.*, 1962, **27**, 4643–4646.

32 V. Zaikin and J. Halket, *A Handbook of Derivatives for Mass Spectrometry*, IM Publications, Chichester, UK, 2009.

33 F. L. Wiseman, *J. Phys. Org. Chem.*, 2012, **25**, 1105–1111.

34 S. P. Asprey, B. W. Wojciechowski, N. M. Rice and A. Dorcas, *Chem. Eng. Sci.*, 1996, **51**, 4681–4692.

35 F. Susanne, D. S. Smith and A. Codina, *Org. Process Res. Dev.*, 2012, **16**, 61–64.

36 J. Clayden, N. Greeves and S. Warren, *Organic Chemistry*, Oxford University Press, Oxford, Second Edi., 2012.

37 *Spectral Database for Organic Compounds (SDBS); NMR spectrum No.: 1224HSP-02-012; http://sdbs.db.aist.go.jp/ (Accessed June 5, 2015)*, .

38 M. Liu, M. A. Little, K. E. Jelfs, J. T. A. Jones, M. Schmidtmann, S. Y. Chong, T. Hasell and A. I. Cooper, *J. Am. Chem. Soc.*, 2014, **136**, 7583–7586.

39 N. Y. Gorobets, S. A. Yermolayev, T. Gurley, A. A. Gurinov, P. M. Tolstoy, I. G. Shenderovich and N. E. Leadbeater, *J. Phys. Org. Chem.*, 2012, **25**, 287–295.

40 D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges and B. D. Sykes, *J. Biomol. NMR*, 1995, **5**, 67–81.

41 *Org. Synth.*, 1995, **72**, 86.

42 *Org. Synth.*, 2000, **77**, 236.

43 N. O'Reilly, N. Giri and S. L. James, *Chem. - A Eur. J.*, 2007, **13**, 3020–3025.

44 N. Giri, M. G. Del Pópolo, G. Melaugh, R. L. Greenaway, K. Rätzke, T. Koschine, L. Pison, M. F. C. Gomes, A. I. Cooper and S. L. James, *Nature*, 2015, **527**, 216–220.

45 J. R. Holst and A. I. Cooper, *Adv. Mater.*, 2010, **22**, 5212–5216.

46 A. I. Cooper, *Angew. Chem. Int. Ed.*, 2012, **51**, 7892–7894.

47 M. Mastalerz, *Chem. - A Eur. J.*, 2012, **18**, 10082–10091.

48 M. E. Briggs, K. E. Jelfs, S. Y. Chong, C. Lester, M. Schmidtmann, D. J. Adams and A. I. Cooper, *Cryst. Growth Des.*, 2013, **13**, 4993–5000.

49 J. T. A. Jones, T. Hasell, X. Wu, J. Bacsa, K. E. Jelfs, M. Schmidtmann, S. Y. Chong, D. J. Adams, A. Trewin, F. Schiffman, F. Cora, B. Slater, A. Steiner, G. M. Day and A. I. Cooper, *Nature*, 2011, **474**, 367–371.

50 M. Mastalerz, *Angew. Chem. Int. Ed.*, 2010, **49**, 5042–5053.

51 T. Hasell, J. L. Culshaw, S. Y. Chong, M. Schmidtmann, M. A. Little, K. E. Jelfs, E. O. Pyzer-Knapp, H. Shepherd, D. J. Adams, G. M. Day and A. I. Cooper, *J. Am. Chem. Soc.*, 2014, **136**, 1438–1448.

52 C. W. Yong, *J. Chem. Inf. Model.*, 2016, **56**, 1405–1409.

53 W. Smith and T. R. Forester, *J. Mol. Graph.*, 1996, **14**, 136–141.

54 T. R. Forester, W. Smith, D. N. Theodorou, R. Kapral, A. J. Hopfinger, W. Pilz and A. J. Zikanova, *J. Chem. Soc. Faraday Trans.*, 1997, **93**, 3249–3257.

55 R. L. Greenaway, D. Holden, E. G. B. Eden, A. Stephenson, C. W. Yong, M. J. Bennison, T. Hasell, M. E. Briggs, S. L. James and A. I. Cooper, *Chem. Sci.*, 2017.

56 M. A. Little, J. Donkin, J. Fisher, M. A. Halcrow, J. Loder and M. J. Hardie, *Angew. Chem. Int.*

Ed., 2012, **51**, 764–6.

57 K. E. Chaffee, H. A. Fogarty, T. Brotin, B. M. Goodson and J.-P. Dutasta, *J. Phys. Chem. A*, 2009, **113**, 13675–84.

58 A. Gafni and Y. Cohen, *J. Org. Chem.*, 1997, **62**, 120–125.

59 R. Wimmer, F. L. Aachmann, K. L. Larsen and S. B. Petersen, *Carbohydr. Res.*, 2002, **337**, 841–849.

60 R. Rymden, J. Carlfors and P. Stilbs, *J. Incl. Phenom.*, 1983, **1**, 159–167.

61 B. Levay, *J. Phys. Chem.*, 1973, **77**, 2118–2121.

62 T. M. Hermans, M. A. C. Broeren, N. Gomopoulos, P. van der Schoot, M. H. P. van Genderen, N. A. J. M. Sommerdijk, G. Fytas and E. W. Meijer, *Nat. Nanotechnol.*, 2009, **4**, 721–726.

Chapter 4. Measuring continuous distributions of size

4.1 Introduction to diffusion NMR

Molar-mass dispersity ($Đ_M$)[1] (Equation 1) is the crucial molecular descriptor used to characterise the properties of any synthetic polymer. It describes the distribution of chain lengths produced during a reaction. By comparing the mass-average molecular weight (M_w), and the number-average molecular weight (M_n), two measures of the polymer mass with different sensitivities to molecular weight, it is possible to obtain the width of this distribution independently of mass.

$$Đ_M = \bar{M}_w / \bar{M}_n \quad (1)$$

Changes in $Đ_M$ are known to have dramatic effects on the bulk properties of polymers in the solid[2–5] and melt[6–8] states, as well as in solution.[9] Both low,[10,11] and high[5,9] dispersities are often targeted to optimise the properties of a system, and dispersity is especially crucial to demonstrate that a given polymerisation system is living.[12]

Dispersities are most commonly determined by size exclusion chromatography (SEC, also referred to as gel-permeation chromatography, GPC).[13–15] Light,[16–18] X-ray,[19–21] and neutron scattering[22–24] are also commonly used, as well as chromatography-coupled,[25–27] or Matrix-assisted laser desorption ionization (MALDI)[28–30] mass spectrometry.

Although SEC has become the standard method of analysis for polymer dispersity measurements,[31,32] situations often arise in which this is either impractical or ineffective. Where dispersities need to be determined *in situ*, such as during a reaction, chromatographic analysis often requires repeated sampling, workup and isolation before analysis. In addition, many poly(esters) and poly(amides) such as poly(ethylene

terephthalate) and nylon, require processing in costly and/or highly toxic solvents such as hexafluoroisopropanol (HFIP).

Although researchers who specialised in poly(esters) and poly(amides) frequently have access to the polymer columns necessary for this research, it can make chromatography prohibitively expensive for those wishing to enter the field. As the majority of research institutions have access to an NMR service, an ability to determine the dispersity of polymers using NMR could broaden access to this field.

Apparent molecular weight distributions can also be affected by the solvent choice,[33,34] which can skew the measured values of M_n and M_w. Thus, where solvent choice is constrained by column compatibility, or polymer solubility, ambiguous values can be obtained.

In this Chapter, we develop a ^{1}H NMR method that allows measurement of dispersity without the need for work up and in any solvent. To avoid polymer entanglement, which can generate misleading results, these measurements are determined in dilute solutions. However, preliminary investigations are performed that measure chain and dispersities *in situ* during a RAFT polymerisation without the need for workup. We acknowledge that this method has been tested *in situ* only twice and on a single styrene RAFT polymerisation. In Chapter 5, we outline a protocol for validating this new method, and establishing its robustness in determining dispersity of polymers in reaction mixtures.

5.1.1 Using NMR to model dispersities: hypothetical data

One possible reason for the success of SEC to date may be that, once calibrated, a RI or UV chromatogram provides a visual guide of the molecular weight distribution.

Mass spectrometry, which also provides this guide, is limited by non-quantitative ionisation effects, which can lead to mass peak integrals of misleading size.[35]

Conversely, light scattering and NMR techniques rely on incorporating experimental data into models, which seek to fit this data to possible values of M_w and M_n.

In the case of NMR, several models exist, which analyse the ways in which polydispersity can be measured from experimental data.[36–38] In cases where a single molecule of definite mass exists in solution, the PFG-NMR response curve is characterised by a straight-line dependence between log signal intensity and g^2 (Figure 4.1a, see also Chapter 2, Section 2.1.1). The diffusion coefficient is thus obtained by differentiating this curve with respect to g.

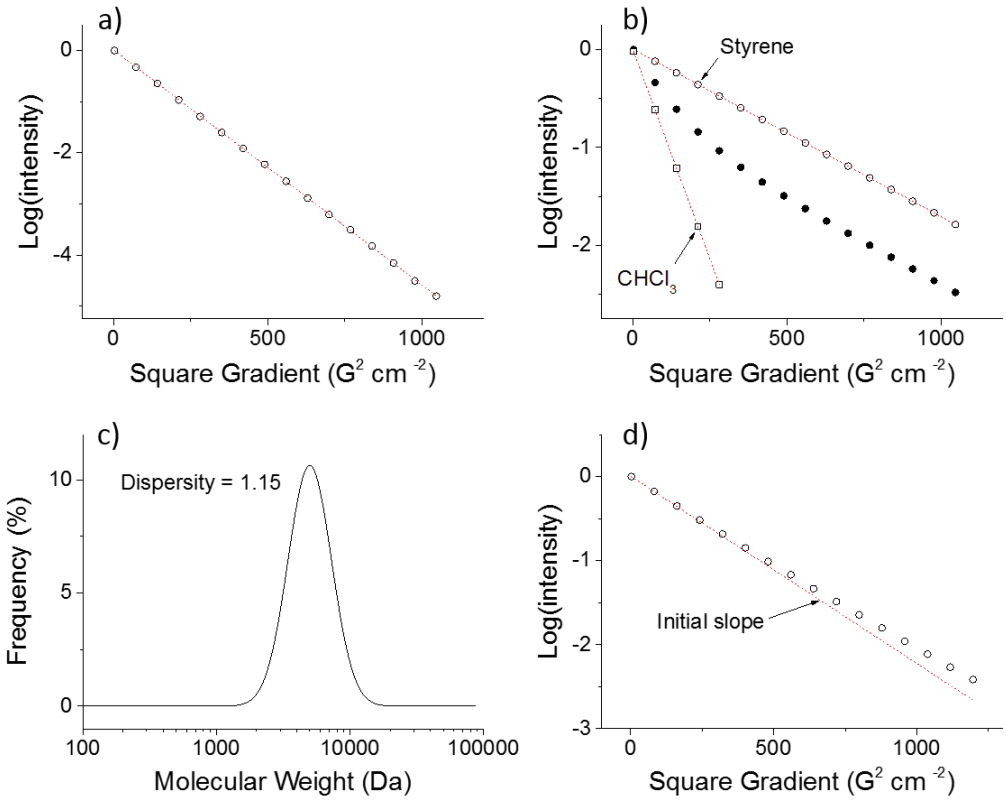


Figure 4.1: (a) ^{1}H PFG-NMR response curve for styrene demonstrates a linear dependence on gradient strength; (b) The PFG-NMR response curve (filled circles) for a two-component (bimodal) system is a product of the individual components. Shown as open symbols are the calculated response curves of styrene (circles) and chloroform (squares), produced by de-convoluting the bimodal data; (c) the molecular weight distribution of polystyrene PS4 (Table 4.1); (d) the response curve observed for PS4 demonstrates a gradual deviation from the initial gradient (red dotted line).

In systems where the two molecules of different sizes are present, but their signals overlap in the ^{1}H NMR spectrum, the intensity can still be analysed, but the response curve is no longer linear. Assuming the two molecules do not interact, the response curve is related to the sum of two independent decays, each relating to one molecule (Equation 3).

$$S = S_1 e^{-bD_1} + S_2 e^{-bD_2} \qquad (3)$$

The response curve takes the form of a 'kinked' line, because it combines two individual decays with very different responses to an increasing gradient field (Figure 4.1b). Where the sizes of two molecules differ by more than about 20 %,[39] the response curves of two contributing molecules can be mathematically disentangled, and D calculated for each.

In solutions that contain polymers, the system comprises molecules that vary continuously in size, but still overlap in the ^{1}H NMR spectrum. Figure 4.1c shows the molecular weight distribution (MWD) for a polymer of $Đ_M$ = 1.5. The PFG-NMR response curve associate with this polymer demonstrates a continuous curve (Figure 1d).[40,41] In principle, curvature is created by summing the decays of each molecule in the series, where the diffusion coefficient of each molecule, D_i, varies systematically (Equation 4).

$$S = S_0 \sum_i e^{bD_i} \qquad (4)$$

Two factors contrive to make this more experimentally complicated. As signal intensity is calculated from peak integrals, larger molecules that have more protons contribute more greatly to the distribution. This accentuates the curvature observed.[42] Counter-acting this affect, because longitudinal relaxation times increase with molecular weight, signal intensities of larger molecules may appear to be decreased under standard experimental conditions.[43–45] However, in the case of polymers, this effect is minimal.[46,47]

Several attempts have been made to approximate the response curve observed using various functions: the Tikhonov regularisation,[36] gamma function,[38] and log/Normal distribution,[37] amongst others.[41,48–50] In each case, the shape of the molecular weight distribution (MWD) is assumed to take some ideal shape (Figure 4.2).

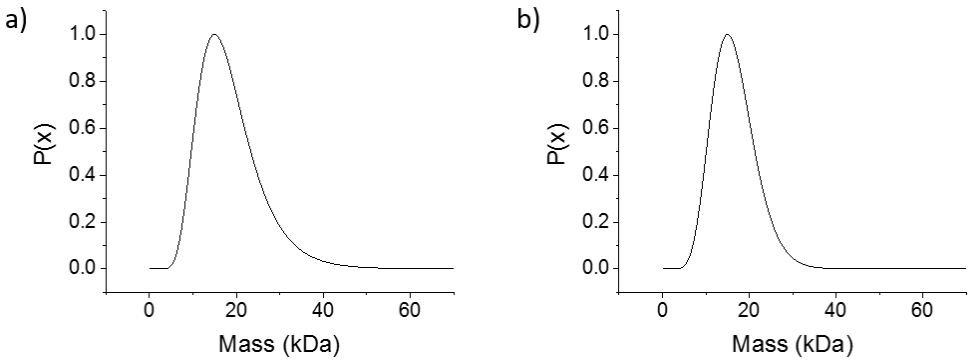


Figure 4.2. (a) Log/Normal and (b) Gamma distributions of hypothetical systems with M_p = 20 kDa and $Đ_M$ = 1.25.

In theoretical investigations, where the studied line shapes are ideal, these can be successful in generating good agreement with known distributions. However, frequently polymers adopt non-ideal distributions; these are prevalent with shorter polymers,[51,52] and with less controlled polymerisations.[53]

Where the shape of the distribution is unknown, these assumptions may limit the maximum accuracy of results, because an ideal line shape is fitted to non-ideal data.

In principle, systematic prediction can be achieved through multi-exponential fitting using programmes such as CONTIN (CONTINuous distributions of exponentials)[41,54] and DECRA (Direct Exponential Curve Resolution Algorithm),[55] or manually *via* a non-linear least-squares fitting.[40] These avoid the problems surrounding assumption of a line shape that may be incorrect.

However, systematic distribution-fitting algorithms attempt to fit a large number of parameters simultaneously.[56] This, combined with relatively high errors inherent in determining the size of peak integrals at high *g*, frequently results in data smoothing. Where improvements to noise can be made, this technique may in the future become invaluable.

One possible solution is to decrease the complexity of the model used to calculate polydispersity to the extent that line-shape is unimportant. In this way, a minimum number of parameters must be fitted, which reduces potential smoothing. Furthermore, because no line shape is assumed, the results are not misleading, but provide a value of $Đ_M$ that can be used to predict the physical properties of a polymer. This solution is described in the following Section, and implemented in Section 3.2.

5.1.2 A simple model for measuring dispersity by PFG-NMR

Non-linearity in the PFG-NMR response curve is related to the polymer MWD in such a way that increased $Đ_M$ is expected to result in a significant rise in curvature (Figure 4.3).

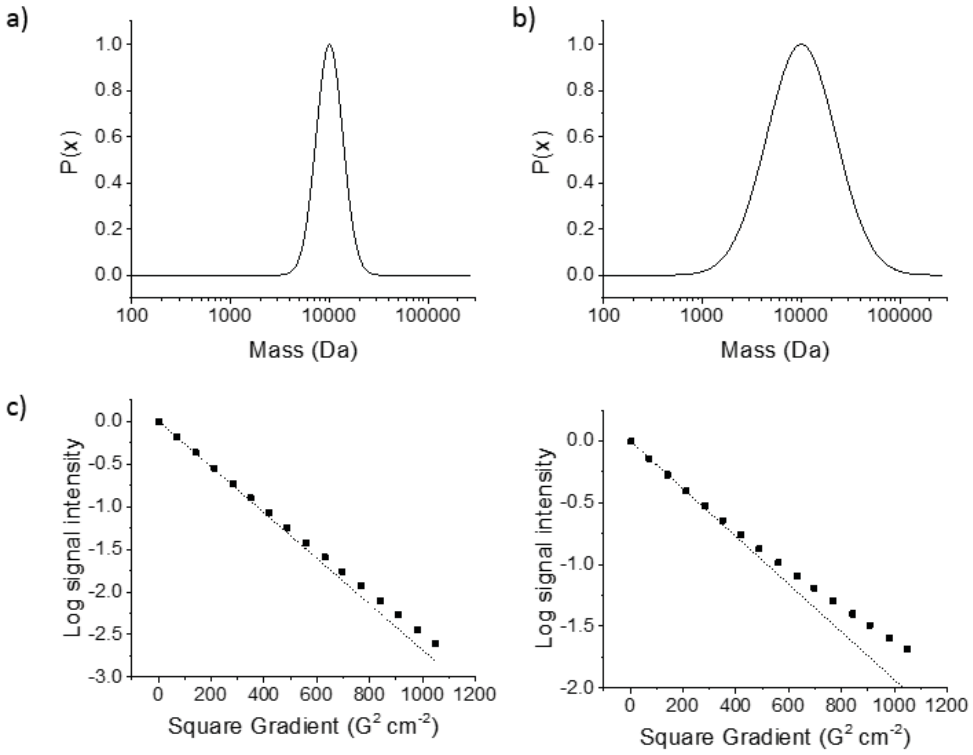


Figure 4.3. Predicted MWDs of two theoretical polymers with $Đ_M$ of (a) 1.1 and (b) 2.0; (c) Calculated PFG-NMR response curves, modelled using Equation 4, representing polymers with $Đ_M$ = 1.1 (left), and with $Đ_M$ = 2.0 (right). In each case, a straight dotted line has been drawn to provide a useful gauge of the curvature of each line.

Data at low *g*, where the gradient is most negative, represent the smallest molecules in the data set. At low *g*, large polymers move so slowly that their signal intensity is unaffected by small changes in the amplitude of the magnetic field. As such, using regression to estimate the initial gradient of the response curve will provide an idea of the size of the smallest molecules in the distribution (Figure 4.4).

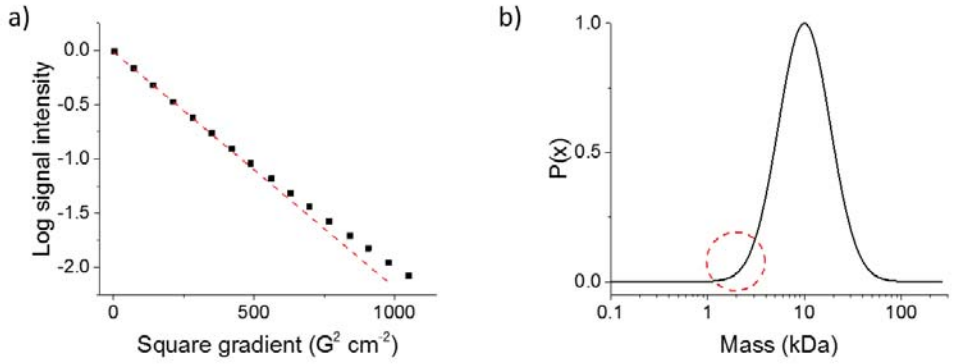


Figure 4.4. (a) PFG-NMR response and (b) MWD of a polymer with $Đ_M$ = 1.5. Theoretical data. Initial gradient is highlighted in (a) by a red dotted line, which represents low-mass molecules, which are highlighted in (b) by a red dotted circle.

Different regression models can be used to estimate the gradient as it changes with *g*. For instance, because Equation 4 suggests that the gradient will increase exponentially, an exponential fit could be used. Noise in this section of the data is relatively low, because signal intensity is high, and signal-to-noise ratio is large. However, errors associated with peak integral calculation are still ~5 %.[39] This can afford a sizeable error in regression, which is amplified by scaling exponentially, and is discussed more in Section 3.2.3.

Furthermore, the shape of the distribution is unknown, and cannot be calculated from PFG-NMR. In cases where the polymer MWD contains a low-mass tail, the initial gradient calculated by exponential regression will differ significantly, even though $Đ_M$ may not (Figure 4.5).

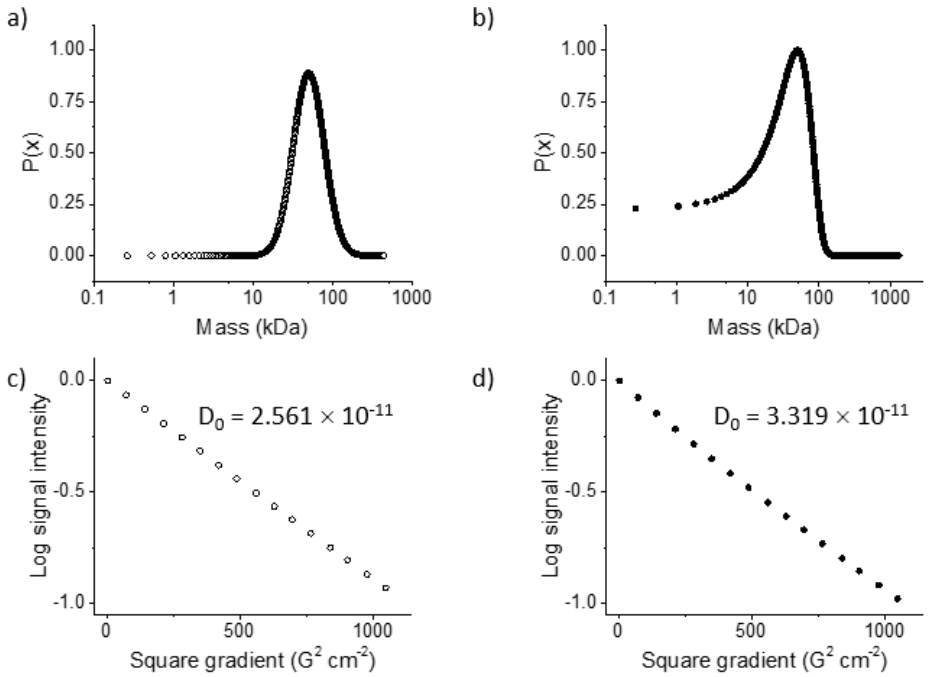


Figure 4.5. Calculated molecular weight distributions for hypothetical polymers with a (a) Log/Normal (b) normal distributions. In each case, $Đ_M$ = 1.25 and M_p = 50 kDa; (c) Calculated PFG-NMR response curves, modelled using Equation 4, representing polymers with Log/Normal (left), and normal (right) distributions. D_0 is obtained from initial gradients, which were calculated by linear regression.

To minimise the effect of noise, and prevent small deviations in the MWD from significantly affecting results, a linear regression can be used, which is relatively invariate to both noise and line shape.

The initial gradient, bD_1, and intercept, S_1, are calculated from the first six data points using linear regression (Figure 4.6a). Using this initial gradient, we calculate a

straight-line ST response curve. Because the PFG response curve of one component has a single gradient and intercept (e.g. styrene, Figure 4.1a), it is convenient to group the initial gradient and intercept as if they correspond to one molecule. For this purpose, we define Component 1 (**C1**, Figure 4.6a, dotted line). This "component" is not real – it does not represent one specific polymer in the mixture. However, it is a useful average of the smallest molecules in solution. By fitting a second component, **C2**, that represents the largest molecules in solution, we can compare the components to determine the spread of sizes in solution and, hence, the dispersity.

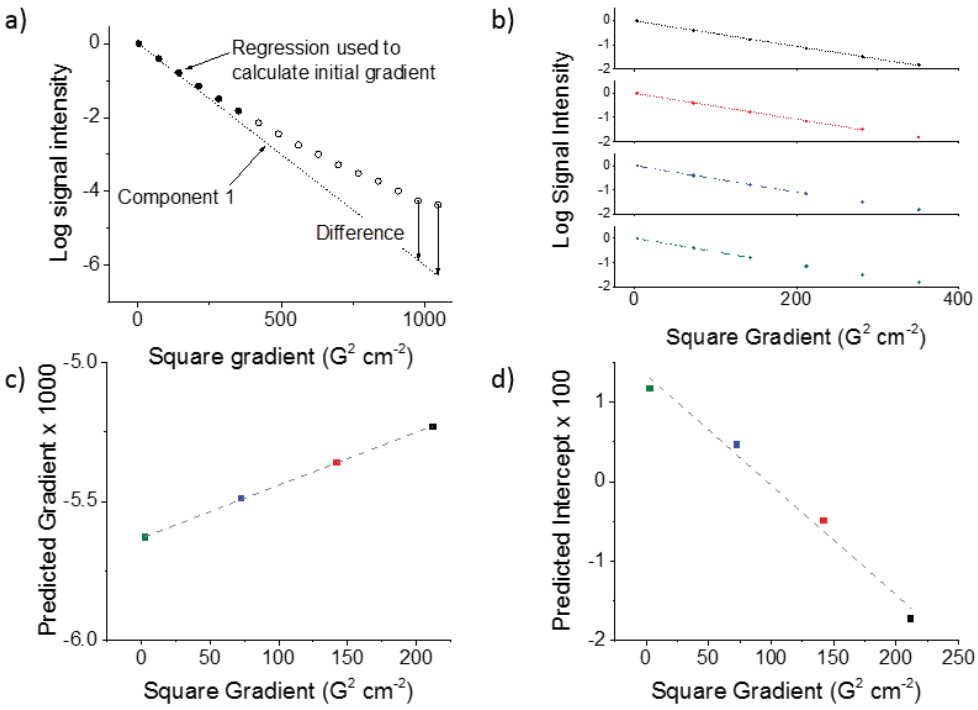


Figure 4.6: (a) Signal decay observed during a PFG-NMR experiment of a hypothetical polymer with $Đ_M$ = 1.40. Data points used to calculate the initial gradient are filled. (b): The gradients and intercept for the first 6 (black), 5 (red), 4 (blue), and 3 (green) data points are calculated; These data are interpolated backwards to estimate the gradient (c) and intercept (d). In each case, data points are colour coded in line with (b). Thus, the intercept calculated using only the first three data points is green, and so on.

As the degree of curvature is related to the breadth of the MWD, calculating the rate at which the PFG-NMR response curve deviates from the initial gradient allows us to quantify the $Đ_M$ without assuming any MWD line shape.

Once the initial gradient has been calculated, this is used to generate a single-component PFG-NMR response curve. The calculated component, **C1**, is characterised by a diffusion coefficient, D_1, and an initial intensity, S_0 (Equation 5).

$$LnC_1 = LnS_0 - bD_1 \qquad (5)$$

where C_1 is the signal intensity associated with component **C1**. By subtracting **C1** from the experimental data, the degree to which the response curve deviates from the initial gradient can be isolated (Figure 4.7).

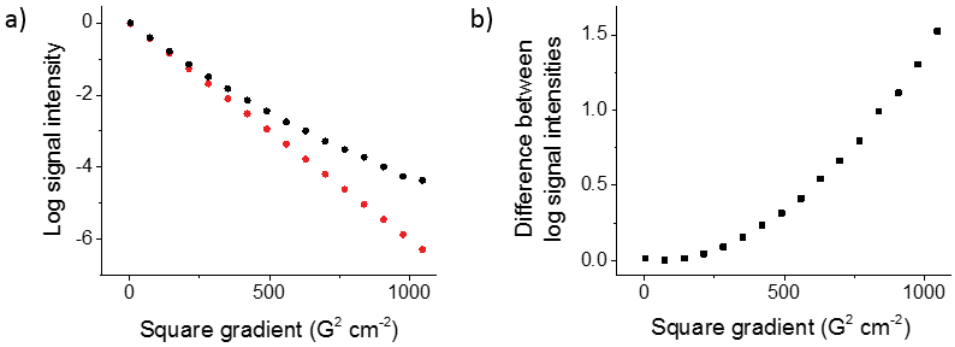


Figure 4.7. (a) PFG-NMR response curve of a polydisperse species (black circles) and the hypothetical PFG-NMR response curve calculated from the initial gradient, C1 (red dotted line); (b) The difference between experimental data and C1.

Classically, attempting to fit polydisperse response curves to two-component PFG-NMR response curves (Equation 3) fails, because four parameters (S_1, S_2, D_1, D_2) are fitted simultaneously. This creates a shallow phase space that is prone to producing suboptimal results. Furthermore, having subtracted **C1** from the experimental data, we

observe that the increasing deviation is non-linear, which confirms that the deviation is not the result of a single additional component.

Here, we assume that the PFG-NMR response curve is composed of two components, **C1** and **C2**, which both take the form of a single exponential ST decay, (Equation 6).

$$S = S_0(Ae^{-bD_1} + Be^{-bD_2}) \quad (6)$$

where A and B are constants that represent overall contributions of **C1** and **C2**, **C2** = bD_2, and D_2 is the diffusion coefficient of the second component.

As **C2** is calculated from data points at high *g*, which affect the initial signal only slightly, S_0 is determined from **C1** only. Expressing Equation 6 logarithmically, and subtracting C1 (Equation 5), we afford Equation 7:

$$LnS_{dev} = \mathrm{Ln}\left(\frac{e^{D_1}}{Ae^{D_1} + Be^{D_2}}\right) \quad (7)$$

where S_{dev} is the signal intensity once the calculated PFG-NMR decay associated with **C1** has been subtracted.

Modelling this non-linear ratio can be accomplished in a number of ways. To determine whether treating this deviation as linear, quadratic or exponential most efficiently enabled the dispersity of polymers to be calculated, it was necessary to test each fitting against a number of polydisperse materials.

4.2 Results and discussion

To test which fitting function was most appropriate, a library of polymers was synthesised for this work (Table 4.1), which was designed to include a range of monomer types. Poly(styrene)s (**PS1–7**), poly(methacryloyloxyethyl phosphoryl-choline)s (**MPC1–2**), poly(vinyl acetate) (**PVA1**), poly(vinyl pyrrolidone) (**PVP1**) and poly(hydroxyethyl acrylate)s (**PHEAA1–3**) were chosen to sample a range of θ

solvents, degree of polymerisation (DP) and functional groups. Where synthesis was conducted by others, this is highlighted in the final column of Table 4.1.

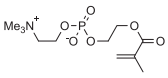
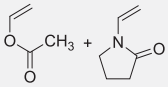
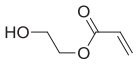
Monomer	DP	M_n[a] / Da	M_w[a] / Da	$Đ_M$[a]	Name
	12	1050	1300	1.24	**PS1**[b]
	14	1300	1500	1.17	**PS2**[b]
	50	3750	4150	1.11	**PS3**[b]
	110	11250	12500	1.12	**PS4**[c]
	136	13950	16350	1.17	**PS5**[c]
	247	25350	28550	1.13	**PS6**[c]
	289	29600	35600	1.20	**PS7**[c]
	136	37950	55550	1.46	**MPC1**[d]
	76	21300	22400	1.05	**MPC2**[d]
	100	8400	11000	1.30	**PVA1**[e]
	50	5600	9700	1.73	**PVP1**[e]
	9, 130	15000	22500	1.50	**PVAP1**[e]
	92	10700	13150	1.23	**PHEAA1**[e]
	164	19050	23400	1.23	**PHEAA2**[e]
	178	20700	24600	1.19	**PHEAA3**[e]

Table 4.1: A library of polymers was synthesised with a variety of chemistries; (a) Molecular weights and dispersities as measured by GPC; (b) GPC characterisation conducted in THF. Polymers synthesised Thanchanok Ratvijitvech at the University of Liverpool; (c) GPC characterisation conducted in THF; (d) GPC characterisation conducted in water, with pH 5 acetic acid buffer. Polymers synthesised by Mike Barrow at the University of Liverpool; (e) GPC characterisation conducted in water, with DMF + 0.1 % LiBr with a Polargel M Column Set at 50°C at a flow rate of 1 mL min^{-1}. Polymers synthesised by Tom Congdon and Ben Martyn at the University of Warwick.

In each case, the ratio created by subtracting **C1** from the experimental PFG-NMR was modelled as a function (F(g), Equation 8).

$$\mathrm{Ln}\left(\frac{e^{D_1}}{Ae^{D_1}+Be^{D_2}}\right) = F(g) \quad (8)$$

This function was assumed to be linear (Equation 9):

$$F(g) \approx Agx + B \quad (9)$$

where x is equal to g^2, and A and B are constants; quadratic (Equation 10):

$$F(g) \approx Ax^2 + Bx + C \quad (10)$$

where C is a constant; or exponential (Equation 11):

$$F(g) \approx \Delta S_0 + Ae^{x/B} \quad (11)$$

where ΔS_0 is the magnitude of the deviation in the absence of diffusion. Using each approach, fitting functions were used to generate the line of best fit with respect to this deviation. The constants calculated for each polymer were compared to $Đ_M$, as calculated by GPC.

4.2.1 Linear fitting function

The simplest treatment of data extracted from the PFG-NMR response is to apply a linear fitting function to the deviation curve. Although the line-shape is non-linear, it is possible to divide the line into two sections, which illustrate the treatment of the data so far. The first six data points, which were used for calculation of the initial gradient, are characterised by an initially small gradient. In some cases, this section of the curve has a slightly negative gradient, resulting from noise in the initial gradient calculations (Figure 4.8).

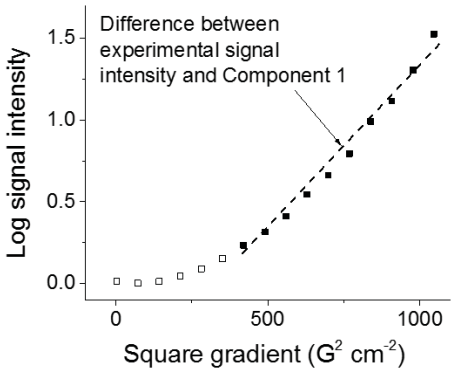


Figure 4.8. PFG-NMR deviation curve. Highlighted, the first six data points (open circles) and the subsequent data points (filled circles). A line of best fit (linear, dotted black line) is drawn through the subsequent data points.

The second portion of the data, above 400 G^2 cm^{-2}, is characterised by a significantly larger gradient, is used to quantify the rate of deviation. This gradient, $\delta(S_{dev})/\delta(g^2)$, was found to be highly dependent upon both b, and the natural logarithm of the signal intensity at 95 % signal intensity (S_{95}, Figure 4.9).

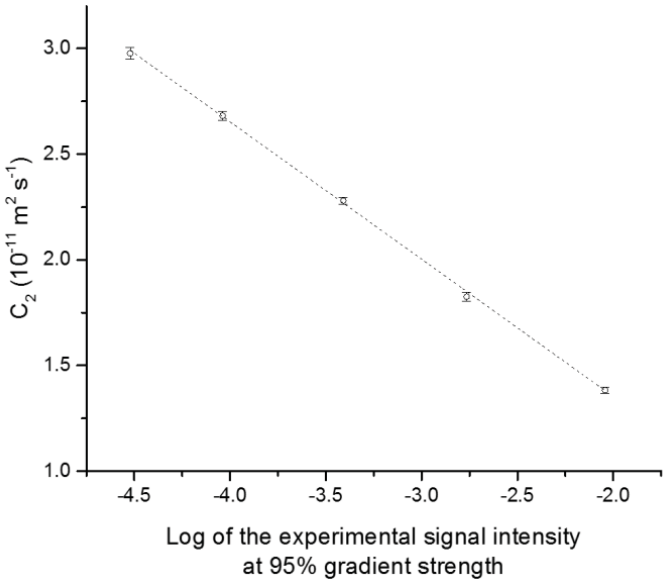


Figure 4.9: The (log) experimental signal intensity at 95 % gradient strength was found to be directly proportional to the gradient of the measured deviation.

The magnitude of $\delta(S_{dev})/\delta(g^2)$ is dependent on S_{95}, as the deviation increases with signal loss. Thus, increasing experimental parameters to increase signal loss will increase the size of the deviation without increasing the size of the gradient. To obtain data that is independent of signal loss, and thus experimental parameters), it is necessary to divide the value of $\delta(S_{dev})/\delta(g^2)$ by the value of S_{95}.

Secondly, it was thought that the dependence of $\delta(S_{dev})/\delta(g^2)$ on the value of the b-factor arose from the non-linearity of the deviation. Thus, the gradient of the deviation increases as Δ and δ are increased. To confirm that this was not a result of a single parameter, PFG-NMR measurements were carried out on **PVA1** at varying magnitudes of Δ and δ, while b was kept constant. No change in $\delta(S_{dev})/\delta(g^2)$ was observed, which was thought to confirm that $\delta(S_{dev})/\delta(g^2)$ was independent of Δ and δ individually.

Thus, by dividing Δ and δ by both the value of the b-factor and $Ln(S_{95})$, a normalised gradient value, $\bar{\delta}$, is produced that is independent of experimental parameters (Equation 12).

$$\bar{\delta} = {}^{(\delta S/\delta g)}\Big/_{(b \cdot LnS_{95})} \qquad (12)$$

Using this above approach, $\bar{\delta}$ was calculated for all polymers in this library. $\bar{\delta}$ was found to agree with the dispersity, as measured by GPC, from 1.05 to 2, (Figure 4). A line of best fit was estimated (Figure 4.10, red dotted line) using the Levenberg–Marquardt algorithm for linear-regression in Origin 2015. Using polymers in Table 4.1 as a calibration library, we relate the polydispersity to the experimental value of $\bar{\delta}$ (Equation 13).

$$\bar{\delta} = 1 - Đ_M{}^{-A} \qquad (13)$$

where A is a scaling parameter, equal to 1.655×10^{-11}.

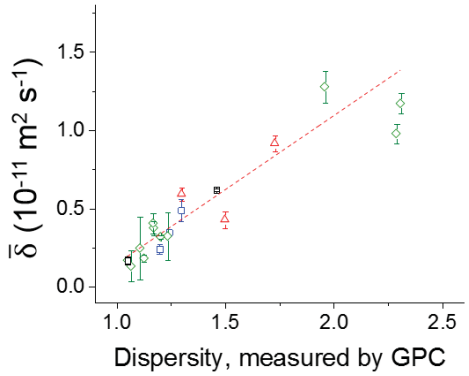


Figure 4.10: $\bar{\delta}$ varies linearly with polydispersity. Polydispersities were measured by GPC. Error bars shown correspond to three standard deviations (3σ). Colours are coordinated by polymer chemistry: MPC1–2 (black circles), PHEAA1–3 (blue squares), PS1–7 and $PS_{Mix}1$–3 (green diamonds), PVA1, PVP1 and PVA/PVP (red triangles).

The signal-to-noise ratio (s/n) of this type of fitting function can be calculated by dividing the overall change in $\bar{\delta}$ between 1 and 2 $Đ_M$ by the average error in readings. In this case, s/n = 20.3, which indicates good confidence can be placed in the magnitude of results, and their relationship with the measured $Đ_M$.

4.2.2 Quadratic fitting function

Whilst data-segmentation was necessary in the case of linear fit, this was not necessary in the case of quadratic fit. In this case, the increasing gradient associated with a quadratic curve reflected the line-shape of the deviation more closely (Figure 4.11).

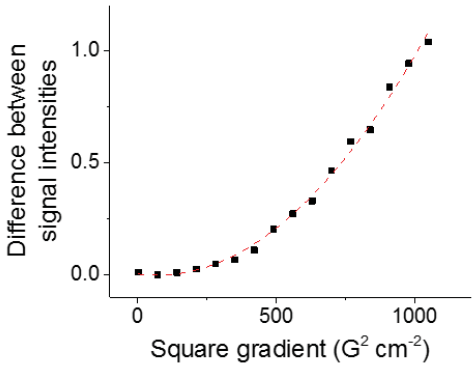


Figure 4.11. PFG-NMR difference curve (black circles) and a quadratic line of best fit (red dotted line).

Parabolic fits were applied to each polymer in the library to test the degree to which the parameters of this fit reflected the dispersity of the material being measured. Again, values were normalised by $S_{95,}$ to remove effects associated with increased signal loss. Once normalised, neither linear nor quadratic coefficients were consistently proportional to the b-factor (Figure 4.12). In the case of **PHEAA1**, a statistically significant positive correlation was noted (Figure 4.13a). However, the change in A observed between b-factor values of 6000 and 11,000 rad^2 s T^{-2} was less than two standard deviations of experimental error. This was also true in the case of **PVP1**, but a negative correlation was noted. In many cases, no correlation was apparent.

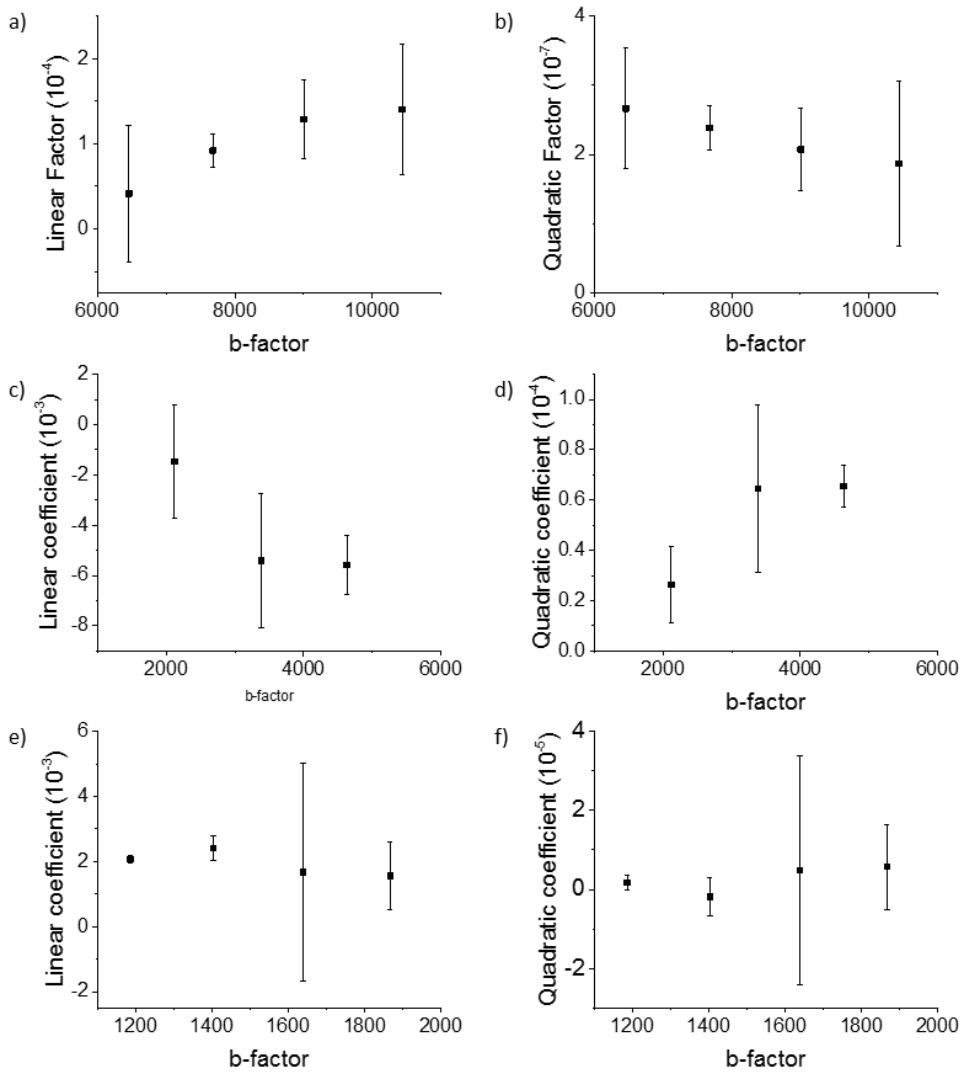


Figure 4.12. Linear (left) and quadratic (right) coefficients for a series of PFG-NMR experiments on PHEAA1 (a,b), PVP1 (c,d) and PS2 (e,f).

Each coefficient was averaged across all experiments. This treatment was repeated for **MPC1–2**, **PS1–3**, **PHEAA1–2** and **PVA1** (Figure 4.13).

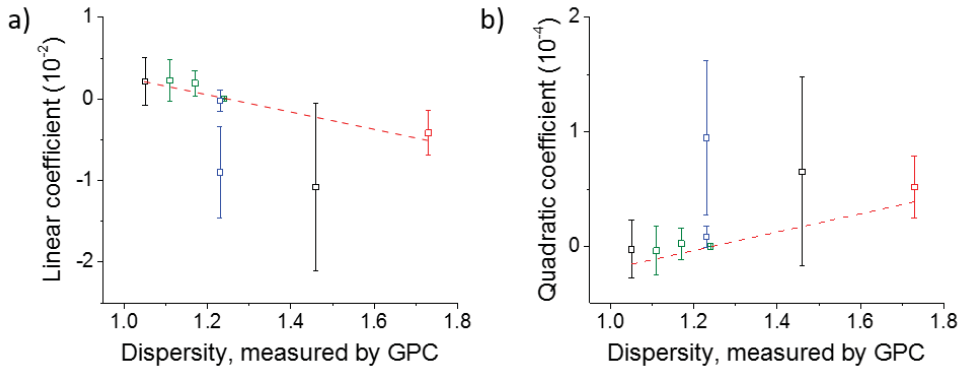


Figure 4.13: (a) Linear and (b) quadratic coefficients of the library of polymers compared to the dispersity measured by GPC. Colours are coordinated by polymer chemistry: MPC (black), PHEAA (blue), PS (green), PVA1, PVP1 and PVP (red).

Trends associated with quadratic fitting were found to be less significant than those obtained by a simple, linear fit. With increasing $Đ_M$, linear coefficients were found to decrease. Between 1 and 2 $Đ_M$, the linear coefficient decreases from 0.0021 to -0.0041 (±0.003) $m^2\ s^{-2}$. Over the same period, the exponential coefficient was found to increase from -2.26×10^{-6} to 5.21×10^{-5} ($\pm\ 3\times10^{-5}$) $m^2\ s^{-2}$.

In each case, the change is greater than the average value of one standard deviation, indicating that it may be a useful parameter for measuring $Đ_M$. However, the opposing changes in A and B may weaken the statistical significance of each.

This may be a product of an attempt to fit an additional coefficient to data with significant experimental noise. However, it is more likely that this represents poor line shape match between the experimental deviation, and the quadratic curve. The decreasing value of the linear coefficient indicates that deviation increases at a rate less than g^2.

One method for determining whether this exponent, whilst less than 2, is equal for all polymers, is to attempt to fit each deviation curve to Equation 11. This attempts to

model an exponential curve, in which the exponent is fitted, in addition to the pre-exponential coefficient.

4.2.3 Exponential fitting function

The deviation in each data set was treated identically, with Equation 11 being parameterised using the Levenberg–Marquardt algorithm used in Origin 2015. In the absence of diffusion, signal intensity is 100 %. Thus, no deviation resulting from additional signal intensity is possible. For this reason, exponential fitting was initially modelled under the constraint $\Delta S_0 = 0$ (Figure 4.14a).

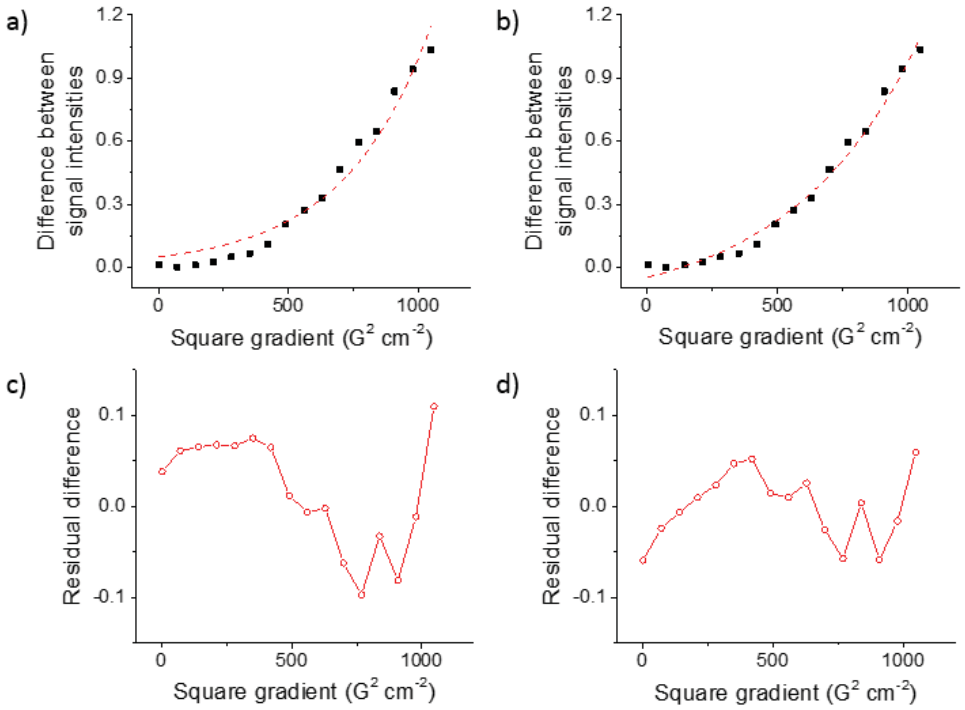


Figure 4.14. PFG-NMR deviation curve (black squares) lines of best fit (red dotted line) were generated by fitting the deviation curve to Equation 11 (a) under the constraint $\Delta S_0 = 0$ and (b) under no constraints; Residual difference between fit and experimental data (red circles). (c) corresponds to (a); (d) corresponds to (b).

Agreement between the experimental and theoretical line shapes appeared to systematically deviate positively at low *g*, and negatively above 600 G^2 cm^{-2}. Conversely, fitting under no constraints generated fits that suggested good representation of line shape. Residual differences were observed to be significantly smaller (Figure 4.14c–d) and less systematic. On this basis, each deviation curve was parameterised with no constraints. However, as ΔS_0 was found to be consistently negative, it was thought to have no significance, in the same way that the intercept in the linear fit was thought to have no physical meaning.

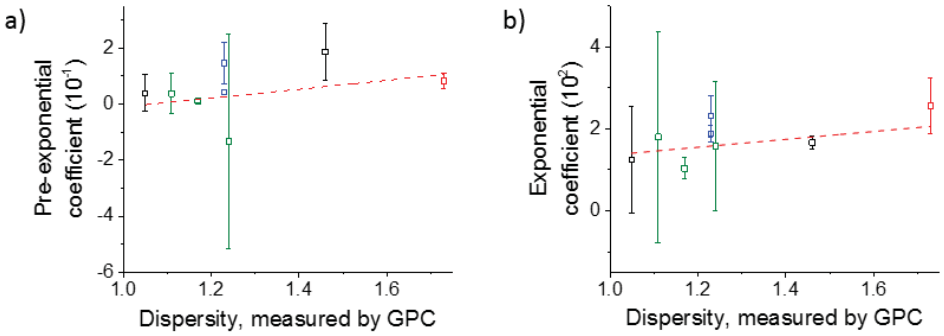


Figure 4.15: (a) Pre-exponential and (b) exponential coefficients extracted from PFG-NMR of the library of polymers compared to the dispersity measured by GPC. Colours are coordinated by polymer chemistry: MPC (black), PHEAA (blue), PS (green), PVA1, PVP1 and PVP (red).

Both pre-exponential coefficient, and exponential coefficient were found to increase between 1 and 2 $Đ_M$ (Figure 4.15). The pre-exponential coefficient was found to increase from 0.039±0.065 to 0.082±0.03 m^2 s^{-1}. B was found to increase from 125±130 to 256±68 m^2 s^{-1}. In each case, the magnitude of the increase was smaller than the average standard deviation associated with experimental error, which indicated that this would not be a reliable method of determining $Đ_M$.

4.2.4 Comparison of fitting functions

The magnitudes of each parameter differed greatly. In the case of the linear fit, as all values were normalised by divining by b, constants derived were of the order 10^{-11}. In the case of quadratic and exponential fits, parameters ranged from 10^{-5} to 10^{2}. To compare the statistical significance of each fitting function, the magnitude of the change experienced in each parameter between D_M = 1 and 2, Δ, was divided by the average standard deviation, σ (Table 4.2).

Function	Parameter	Δ/σ
Linear	A	20.3
Quadratic	A	3.15
	B	2.60
Exponential	A	1.72
	B	1.06

Table 4.2. The significance of each parameter, calculated by dividing the change between 1 and 2 D_M by the average standard deviation.

Significance was noted to increase in the order of exponential < quadratic < linear. In the case of both exponential and quadratic fitting functions, this was generally a result of the fitted curve out-pacing increases in F(g). This indicated that, on average, F(*g*) may increase at a rate that was smaller than x^2, and significantly smaller than e^x.

Unreliable correlation of B (Equation 10) with the value of the b-factor indicated that the rate at which F(*g*) increased was not consistently larger than quadratic. This inconsistency resulted in poor overall fittings and relatively high s/n. By decreasing

the number of variables fitted, and sampling only one half of the F(*g*) curve, linear fitting facilitated a significantly more reliable analysis. This resulted in an almost eight times greater significance than quadratic fittings, and a greater than ten times increase over the significance of the exponential factor.

4.3 Tracking polymerisations by diffusion NMR.

To probe the synthetic usefulness of this technique, ^{1}H PFG-NMR was used to measure the dispersity of a styrene polymerisation carried out using reversible addition− fragmentation chain-transfer (RAFT) polymerization over 240 hours.[57] In the absence of UV radiation, this reaction is dependent upon continuous heat, and, at room temperature, the reaction slows significantly.[58]

A 5 mm NMR tube with a J Young's tap was used to exclude oxygen, and the sample was degassed with nitrogen prior to being sealed. 4,4′-azobis(4-cyanovaleric acid) (ACVA) was used as initiator, and 4-cyano-4-[(dodecyl-sulfanylthiocarbonyl) sulfanyl]pentanoic acid was used as chain transfer agent (CTA). A styrene : ACVA : CTA ratio of 50:1:2 was used to target a DP of 50, at a styrene concentration of 0.25 g mL^{-1}.

The reaction was carried out at 100 °C in 1,4-dioxane-d_8. The sealed NMR tube was heated for 12 hours, after which it was cooled to 25 °C to facilitate NMR analysis. This process was repeated for 60 hours. Three ^{1}H PFG-NMR experiments were conducted at each time point, taking roughly 1 hour per measurement. After five iterations, the reaction was heated continuously for 180 hours, and a final series of ^{1}H NMR experiments was run after cooling.

As expected,[59] dispersity was observed to decrease with increasing conversion (Figure 4.16a). At 240 hours, and 61 % conversion, $Đ_{NMR}$ was measured to be 1.11.

The mixture was subsequently quenched into ethanol and the polymer was isolated by filtration. $Đ_{GPC}$ (measured in THF) was found to be 1.09, in close agreement with the value determined from $\bar{\delta}$.

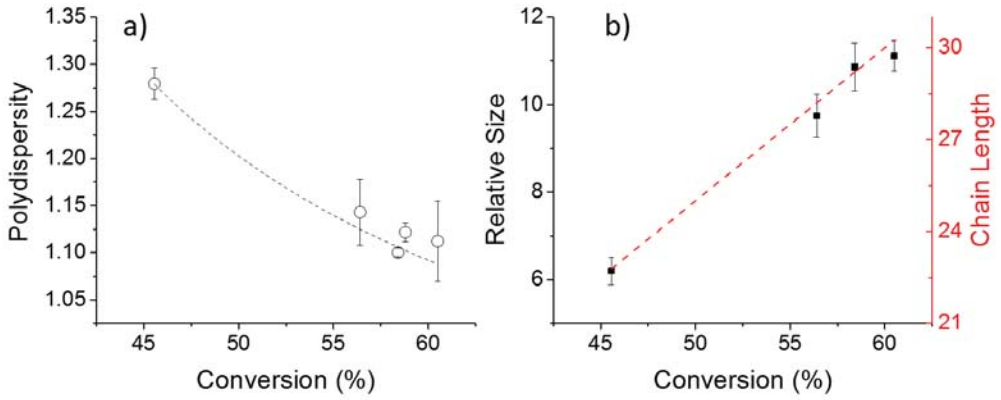


Figure 4.16: (a) Polydispersities (open circles), calculated from ^{1}H PFG-NMR, for a RAFT polymerisation of styrene over 240 h. Dotted line represents a line of best fit; (b) The size of the polymer relative to styrene (black squares) increases with conversion, but is significantly smaller than the calculated chain length (red dotted line).

The conversion was measured by comparing the integral area associated with the SC*H*$_2$ protons of the CTA (3.29 ppm) with the integral area of the poly(styrene) found between 6.5 and 7.2 ppm in the ^{1}H NMR spectrum.[60] From this, we calculated the average chain length, which was compared to the relative size of the polymer. Relative size was calculated by comparing the diffusion coefficients of poly(styrene) and remaining unreacted styrene within the same sample.

Using this comparison, rather than an absolute measurement of the poly(styrene) diffusion coefficient corrected for changes in viscosity associated with increasing polymer chain length.[61] Absolute measurement of M_n by PFG-NMR at concentrations much greater than 5.0 % (w/w) can be problematic, as is discussed in detail below, and was not attempted.

Because polymer motion is segmented,[62,63] allowing the polymer to 'fold in' on itself, the solvodynamic radius (R_s) increases with chain length, but is found to increase much less quickly (Figure 4.16b). The diffusion coefficient (inversely proportional to R_s) can be related to the number of monomers in the chain (N) by a power law (Equation 9).[63]

$$D = \alpha N^{-\beta} \qquad (8)$$

where α and β are experimentally-determined scaling parameters. These parameters are dependent upon the solubility of the polymer in the solvent chosen.[64,65] Thus, individual calibrations must be done for each polymer/solvent mixture, although similarities have been noted between systems where polymers are in theta solvents.[66]

Using **PS1–7**, scaling parameters for poly(styrene)s in chloroform, and was calibrated, allowing for determination of M_n by PFG-NMR (Figure 4.17).

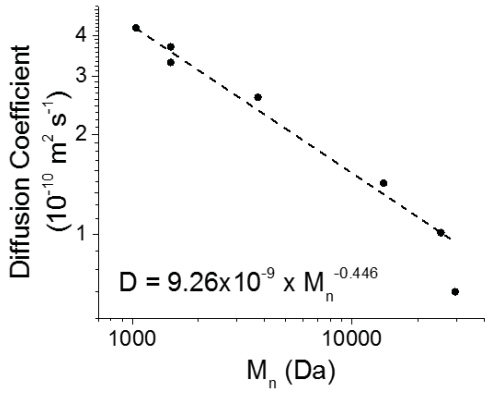


Figure 4.17. Diffusion coefficients of PS1–7 compared with molecular weight (M_n). The line of best fit is represented by a dotted black line. The equation of this line allows M_n to be calculated from diffusion coefficients.

Where concentrations rise above a critical concentration, and polymer chains begin to interact, the measured diffusion coefficients no longer correlate with the molecular

weight of the polymer.[67] Instead, they correlate with the molecular weight of aggregates, which may appear significantly larger.

Concentrations at which polymer interactions were found to interfere with calculations of both M_n and $Đ_{NMR}$ were found to vary between polymers. However, concentrations of 5 % (w/w) were found to consistently remove any interaction effects. Thus, with the exception of the *in situ* RAFT polymerisation, all measurements in this work were conducted at concentrations of 5.0 % (w/w).

In spite of this, $Đ_{NMR}$ for *in situ* polymerisation were found to agree well with $Đ_{GPC}$, indicating that polystyrene chains in dioxane solutions did not entangle significantly.

4.4 High Đ and non-ideal line shapes

To investigate the flexibility of this technique to multimodal and non-ideal line shapes, a series of well-defined polymers (**PS4–7**) were blended, to afford three mixed samples (**PS_{Mix}1–3**) with programmable line shapes (Figure 4.18a–c). Average molecular weights and dispersities were calculated by GPC for each mixture and compared with those calculated by NMR (Figure 4.18d).

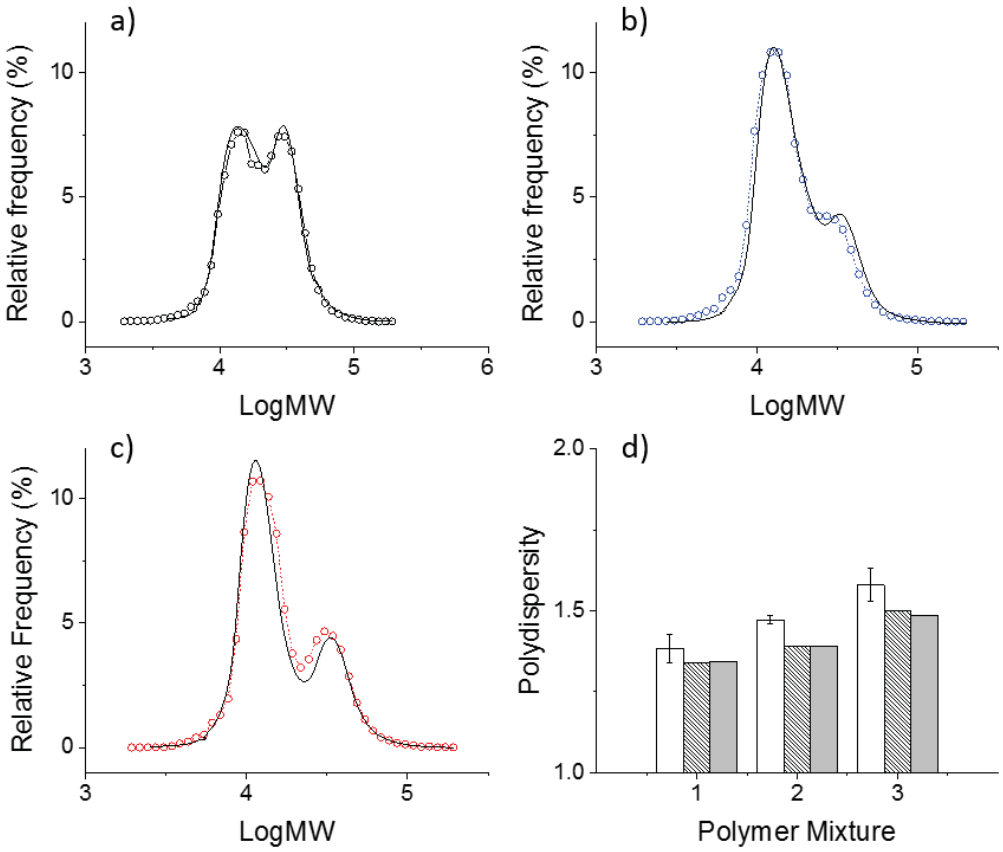


Figure 4.18: Three polymer mixtures were synthesised by mixing PS4–7 in different ratios. (a) PS_{Mix}-1; (b) PS_{Mix}-2; (c) PS_{Mix}-3. The predicted MWD of each polymer mix shown as open circles with dotted lines. GPC data is shown as a solid black line, and was measured in THF; (d) a comparison between polydispersities as measured by NMR (open), GPC (hatched) and predicted (grey).

In all cases, good agreement was found between the NMR-measured polydispersities and those measured by GPC. Slight over-estimation of polydispersities by ^{1}H NMR was noted, which may reflect the amplification of signal at high M_n previously discussed, stemming from the larger number of protons in longer polymer chains.

To investigate whether this overestimation may be the result of solvent choice, diffusion coefficients and NMR-measured dispersities of **PS_{Mix}1–3** were measured in THF-d_8. No statistically significant differences in NMR-measured dispersity were

measured, which indicated that solvent choice had not played a part. Instead, it was thought that the amplification of high-mass molecular weight signals in PFG-NMR response curves (see Section 3.1) was responsible for skewing the weight of results. Because each of the measured polymers had a non-ideal line-shape that was significantly broader, a higher proportion of high-mass polymer species were present, which had not been calibrated in the initial linear model.

Finally, to test the applicability of this technique for measuring the polydispersity of high performance polymers, a broad polymer standard of poly(ethylene terephthalate) (PET) with a GPC-measured polydispersity of 1.89 was dissolved in HFIP at a concentration of 5.0 % (w/w). A lock tube containing $CDCl_3$ was placed inside the 5 mm NMR tube to enable deuterium lock to be carried out. The same process as described above was carried out, affording an NMR-measured polydispersity of 1.67 $\pm$ 0.16. The error was calculated from the average standard deviation associated with experimental error for the linear fit.

This difference appeared to be consistent with the errors associated with both NMR-measured dispersity, and GPC measured dispersity. To confirm that this difference was not related to the presence of $KOO(C)CF_3$ buffer in the GPC measurements, diffusion measurements were repeated in the presence of $KOO(C)CF_3$, although no statistically significant changes in data were noted.

4.5 Conclusions.

In conclusion, we have demonstrated a simple and accessible technique which allows automated analysis of disperse mixtures by ^1H PFG-NMR. The utility for this technique in measuring the dispersities in a wide range of polymer chemistries has been shown, as well as its utility in accurately measuring $Đ_M$ *in situ* during a reaction.

We have demonstrated the tolerance of this technique to bimodal and non-ideal line shapes that may be encountered during reaction, and in cases of polymer blends. This tolerance should be explored systematically, and a protocol for generalising this technique is described in Chapter 5.

As most chemistry departments have access to an NMR spectrometer with a z-gradient coil, this is a widely available and highly accessible technique. The applicability of this technique, specifically to high-performance polymers, where GPC analysis can be prohibitively expensive, broadens its scope.

Many examples exist where proteins form mixtures of monomers, dimers, and oligomers in solution. Currently, diffusion NMR is used to identify the average size of the aggregates formed, although this is usually combined with computationally-generated models, or crystal structures. We envision, however, that detailed analysis of the diffusion response curves could afford information not only of the size, but also the spread of structures.

4.6 Experimental.

All reagents were purchased from Sigma Aldrich. Styrene was passed through basic alumina to remove the inhibitor before use. All other chemicals were used as received. Several polymers were provided by collaborators, and these details are listed in Table 4.1. Details of PFG-NMR are described in Chapter 2.

GPC: PS1–3 were synthesised by Thanchanok Ratvijitvech. All polystyrenes were analysed on a Viscotek system, which comprised a GPC_{max} solvent/sample module and a TDA302 module housing refractive index, viscometer and light-scattering detectors. The system was fitted with 2× ViscoGel HHR-H columns housed in an oven at 40 °C, and THF was used as the eluent. MPCs were synthesised by Michael Barrow and analysed on the same system, using aqueous eluent buffered to pH 5.0 using acetic

acid. For **PVA1**, **PVP1** and **PVAP1** were synthesised by Tom Congdon at the University of Warwick. **PHEAA1–3** were synthesised by Ben Martyn at the University of Warwick. All "Warwick samples" were analysed in water, with DMF + 0.1 % LiBr with a Polargel M Column Set at 50°C at a flow rate of 1 mL min^{-1}. Systems were calibrated with low molecular weight polystyrene standards, and the OmniSEC Universal Calibration method was used.

Synthesis of polystyrene. Linear polystyrenes were synthesised by RAFT polymerisation.[57] In an example synthesis, styrene (3.12 g, 30.0 mmol, 100 eq.) was added to a solution of ACVA (42 mg, 0.15 mmol, 0.5 eq.) and 4-cyano-4-[(dodecylsulfanylthiocarbonyl)sulfanyl]pentanoic acid (121 mg, 0.3 mmol, 1 eq.) in dioxane (12.5 mL). The solution was deoxygenated by nitrogen purging for 30 minutes. The reaction was then heated to 90 °C in a sealed Schlenk tube for 90 hours. After 90 hours, the reaction was cooled to room temperature and the product precipitated into ethanol. Solids were isolated by filtration and dried *in vacuo* at 40 °C for 16 hours to afford **PS4** as a white solid in an 89 % yield. **^{1}H NMR** (CDCl3, 400 MHz) δ = 0.88 (m, 3H), 1.05–2.50 (m, 596H), 3.25 (m, 2H), 6.20–7.25 (m, 797H) ppm. **IR** 3059, 2922, 1712, 756, 697 cm^{-1}. **GPC** M_n = 13,970, M_w = 16,360, PDI = 1.17. For other polymers, the molecular weight of the polystyrene was varied by adjusting the stoichiometric quantity of RAFT agent and initiator.

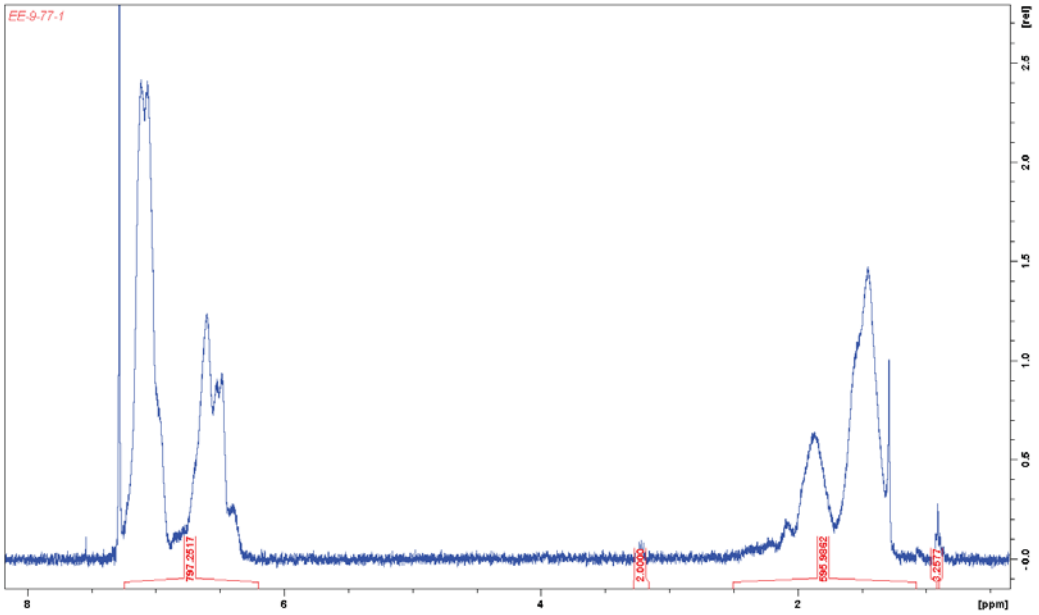


Figure 4.19. ^{1}H NMR spectrum of PS4.

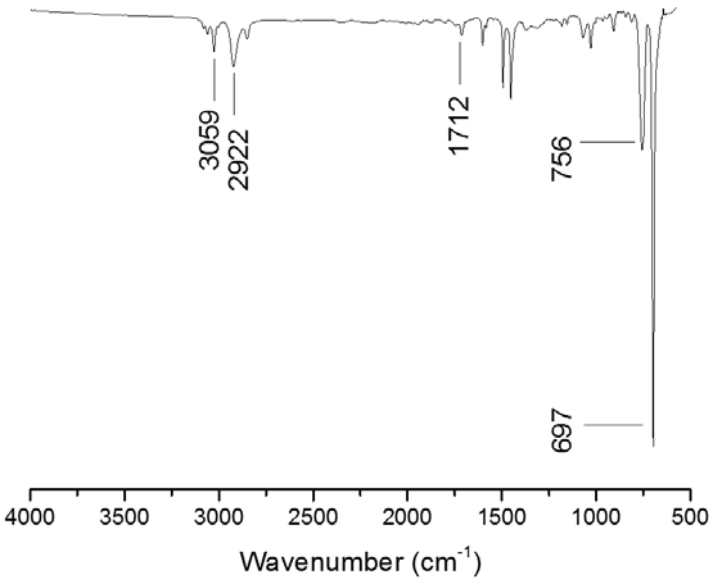


Figure 4.20. Infrared spectrum of PS4.

4.7 Bibliography.

1 R. F. T. Stepto, *Pure Appl. Chem.*, 2009, **81**, 351–353.

2 D. Broseta, G. H. Fredrickson, E. Helfand and L. Leibler, *Macromolecules*, 1990, **23**, 132–139.

3 M. J. Struglinski and W. W. Graessley, *Macromolecules*, 1985, **18**, 2630–2643.

4 D. M. Cooke and A.-C. Shi, *Macromolecules*, 2006, **39**, 6661–6671.

5 L. Lu, T. Zheng, T. Xu, D. Zhao and L. Yu, *Chem. Mater.*, 2015, **27**, 537–543.

6 S. M. Sabzevari, I. Cohen and P. M. Wood-Adams, *Macromolecules*, 2014, **47**, 8033–8040.

7 S. M. Sabzevari, I. Cohen and P. M. Wood-Adams, *Macromolecules*, 2014, **47**, 3154–3160.

8 F. Bueche, *J. Polym. Sci.*, 1960, **43**, 527–530.

9 D. L. Huntson and M. M. Reischman, *Phys. Fluids*, 1975, **18**, 1626.

10 W. H. Blackburn and L. A. Lyon, *Colloid Polym. Sci.*, 2008, **286**, 563–569.

11 J. Listak, W. Jakubowski, L. Mueller, A. Plichta, K. Matyjaszewski and M. R. Bockstaller, *Macromolecules*, 2008, **41**, 5919–5927.

12 G. Moad and D. H. Solomon, *The Chemistry of Radical Polymerization (2)*, Elsevier, Amsterdam, GB, 2005.

13 G. H. Lathe and C. R. J. Ruthven, *Biochem. J.*, 1956, **62**, 665–674.

14 A. M. Striegel, W. W. Yau, J. J. Kirkland and D. D. Bly, *Modern Size-Exclusion Liquid Chromatography*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2009.

15 J. C. Moore, *J. Polym. Sci. Part A Gen. Pap.*, 1964, **2**, 835–843.

16 S. Wiegand and W. Köhler, *Molecular Characterization and Analysis of Polymers*, Elsevier, 2008, vol. 53.

17 D. E. Koppel, *J. Chem. Phys.*, 1972, **57**, 4814.

18 R. Alexander-Katz, in *Handb. Polym. Synth., Charact., Process.*, John Wiley & Sons, Inc., 2013, pp. 367–389.

19 M. Mulato and I. Chambouleyron, *J. Appl. Crystallogr.*, 1996, **29**, 29–36.

20 T. Rieker, A. Hanprasopwattana, A. Datye and P. Hubbard, *Langmuir*, 1999, **15**, 638–641.

21 A. Agbabiaka, M. Wiltfong and C. Park, *J. Nanoparticles*, 2013, **2013**, 1–11.

22 P. Bartlett and R. H. Ottewill, *J. Chem. Phys.*, 1992, **96**, 3306.

23 S. R. Aragón and R. Pecora, *J. Chem. Phys.*, 1976, **64**, 2395.

24 J. S. Higgins and H. C. Benoit, *Polymers and neutron scattering*, Clarendon Press, Oxford,

1994.

25 J. Falkenhagen and S. Weidner, in *Mass Spectrometry in Polymer Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2012, pp. 209–235.

26 T. Gruendling, S. Weidner, J. Falkenhagen and C. Barner-Kowollik, *Polym. Chem.*, 2010, **1**, 599.

27 L. Prokai and W. J. Simonsick, *Rapid Commun. Mass Spectrom.*, 1993, **7**, 853–856.

28 M. W. F. Nielen, *Mass Spectrom. Rev.*, 1999, **18**, 309–344.

29 C. A. Jackson and W. J. Simonsick, *Curr. Opin. Solid State Mater. Sci.*, 1997, **2**, 661–667.

30 L. Li, in *MALDI MS*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013, pp. 313–365.

31 D. Held and P. Kilz, in *Monitoring Polymerization Reactions*, John Wiley & Sons, Hoboken, NJ, 2014, pp. 171–199.

32 M. G. Neira-Velázquez, M. T. Rodríguez-Hernández, E. Hernández-Hernández and A. R. Y. Ruiz-Martínez, in *Handbook of Polymer Synthesis, Characterization, and Processing*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2013, pp. 355–366.

33 T. H. Mourey and T. G. Bryan, *J. Chromatogr. A*, 2002, **964**, 169–178.

34 J. G. Bergmann, L. J. Duffy and R. B. Stevenson, *Anal. Chem.*, 1971, **43**, 131–133.

35 V. Zaikin and J. Halket, *A Handbook of Derivatives for Mass Spectrometry*, IM Publications, Chichester, UK, 2009.

36 I. J. Day, *J. Magn. Reson.*, 2011, **211**, 178–85.

37 X. Gong, E. W. Hansen and Q. Chen, *Macromol. Chem. Phys.*, 2011, **212**, 1007–1015.

38 M. Röding, D. Bernin, J. Jonasson, A. Särkkä, D. Topgaard, M. Rudemo and M. Nydén, *J. Magn. Reson.*, 2012, **222**, 105–111.

39 M. A. Connell, P. J. Bowyer, P. Adam Bone, A. L. Davis, A. G. Swanson, M. Nilsson and G. A. Morris, *J. Magn. Reson.*, 2009, **198**, 121–31.

40 B. Håkansson, M. Nydén and O. Söderman, *Colloid Polym. Sci.*, 2000, **278**, 399–405.

41 A. Chen, D. Wu and C. S. Johnson, *J. Am. Chem. Soc.*, 1995, **117**, 7965–7970.

42 P. T. Callaghan and D. N. Pinder, *Macromolecules*, 1985, **18**, 373–379.

43 B. J. Bauer, L. J. Fetters, W. W. Graessley, N. Hadjichristidis and G. F. Quack, *Macromolecules*, 1989, **22**, 2337–2347.

44 P. G. De Gennes, *Macromolecules*, 1976, **9**, 587–593.

45 P. G. De Gennes, *Macromolecules*, 1976, **9**, 594–598.

46 K. Huber, S. Bantle, W. Burchard and L. J. Fetters, *Macromolecules*, 1986, **19**, 1404–1411.

47 W. Hess, *Macromolecules*, 1986, **19**, 1395–1404.

48 E. D. von Meerwall, *J. Magn. Reson.*, 1982, **50**, 409–416.

49 K. F. Morris, B. J. Cutak, A. M. Dixon and C. K. Larive, *Anal. Chem.*, 1999, **71**, 5315–5321.

50 A. Jerschow and N. Müller, *Macromolecules*, 1998, **31**, 6573–6578.

51 C. C. Chen, G. Kaszas, J. E. Puskas and J. P. Kennedy, *Polym. Bull.*, 1989, **22**, 463–470.

52 T. Fukuda, *J. Polym. Sci. Part A Polym. Chem.*, 2004, **42**, 4743–4755.

53 M. Stenzel-Rosenbaum, T. P. Davis, V. Chen and A. G. Fane, *J. Polym. Sci. Part A Polym. Chem.*, 2001, **39**, 2777–2783.

54 S. W. Provencher, *Comput. Phys. Commun.*, 1982, **27**, 229–242.

55 W. Windig and B. Antalek, *Chemom. Intell. Lab. Syst.*, 1997, **37**, 241–254.

56 C. S. Johnson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1999, **34**, 203–256.

57 G. Moad, Y. K. Chong, A. Postma, E. Rizzardo and S. H. Thang, *Polymer*, 2005, **46**, 8458–8468.

58 A. Goto, K. Sato, Y. Tsujii, T. Fukuda, G. Moad, E. Rizzardo and S. H. Thang, *Macromolecules*, 2001, **34**, 402–408.

59 E. Mastan and S. Zhu, *Macromolecules*, 2015, **48**, 6440–6449.

60 S. Perrier and D. M. Haddleton, in *In Situ Spectroscopy of Monomer and Polymer Synthesis*, Springer US, Boston, MA, 2003, pp. 125–146.

61 N. G. McCrum, C. P. Buckley and C. B. Bucknall, *Principles of polymer engineering*, Oxford University Press, Oxford, Second edi., 1988.

62 P. E. Rouse, *J. Chem. Phys.*, 1953, **21**, 1272.

63 C. R. H. Rubinstein Michael, *Polymer Physics*, Oxford University Press, 2003.

64 L. A. Errede, *Macromolecules*, 1986, **19**, 1522–1525.

65 R. Makitra, H. Midyana, R. Prystansky, Y. Vasyutin and L. Bazylyak, *J. Phys. Org. Chem.*, 2005, **18**, 825–832.

66 W. Li, H. Chung, C. Daeffler, J. A. Johnson and R. H. Grubbs, *Macromolecules*, 2012, **45**, 9595–9603.

67 P. T. Callaghan and D. N. Pinder, *Macromolecules*, 1983, **16**, 968–973.

Chapter 5. Conclusions and Future Work

The aim of this thesis was to explore the possibilities of diffusion NMR in characterising solution-state species, by developing methods for interpreting NMR data. Initial experimentation demonstrated that the shape of covalent cages, as well as their size, was an important factor in determining the magnitude of the diffusion coefficient measured.

5.1 Diffusion NMR of covalent cages

Diffusion NMR is commonly used to confirm the structure of large molecular species in solution. However, PFG-NMR measurements of macromolecular species are rarely published alongside data to confirm the quality of this structural assignment. With this in mind, we have developed several key molecular descriptors that can be measured by PFG-NMR, which allow the quality of structural assignment to be gauged from only a small number of experiments.

By changing the solvent used to measure the diffusion coefficients of covalent cages, we have probed the incurvature associated with certain topologies of covalent cages. By increasing the size of solvent molecules in which covalent cages were dissolved, we demonstrated that the incurvature of cage surfaces could be estimated. The diffusion coefficient of **CC7** demonstrated a much greater solvent-dependence, increasing by 17 % from chloroform-*d* to dichloromethane-d_2. **CC3**, a smaller covalent cage, demonstrates a smaller increase of only 7 %. In further work, a larger range of cages would need to be studied to quantify this effect. This effect could be used to determine the topology of cages, as larger cages frequently have more incurvate surfaces than those with smaller numbers of components.

Furthermore, by comparing the diffusion coefficient measured by PFG-NMR with sizes extracted from SCXRD or computationally-generated structures, we have developed a molecular descriptor ρ_r, which describes the anisotropy associated with the external surface of a cage. ρ_r was observed to decrease with Φ_S, and is indicative of molecular shape. In an initial screen, values of ρ_r larger than 0.7 were found to be indicative of anisotropic, short prismatic cages, a ρ_r value between 0.2 and 0.35 was associated with larger, more isotropic structures. **Astetrapod** demonstrated a higher than expected ρ_r value of 0.54, which we rationalised to be the result of large "apexes" that protrude from the centre of the structure. These findings underscore the importance of considering the structural shape of molecules when justifying diffusion data, and demonstrate the differences in diffusion coefficients that can be experienced both between and within cage topologies. In a subsequent high-throughput screen, a wide variety of structural shapes were encountered, which generated a range of ρ_r values. In future work, we hope to expanded and parameterise this technique for an extended range of cage topologies and molecular sizes. We envisage that this could be used to assess the accuracy of PFG-NMR analysis when it is used to confirm the structure of a covalent cage, which has previously not been possible.

5.2 Towards the synthesis of porous monomers

Finally, we synthesised five novel dodecaamide cage molecules, four of which incorporated functional groups that could be further functionalised *via* further reaction. We successfully controlled the kinetics of acylation reactions, synthesising mixtures of functionalised and non-functionalised cages that contained predominantly singly functionalised material. We have purified a desymmetrised cage $\mathbf{1^1 3^5}$, which incorporates a single ethylenediamine vertex, and five cyclohexanediamine vertices in

a [4+6] cage. We have attempted to selectively functionalise these materials *via* acylation reactions. However, attempts to purify acylated mixtures by crystallisation were hampered by the high solubility of desymmetrised cages, and the polarity of remaining amine groups made chromatographic separation impossible. Reactions between reduced cages and symmetrical ketones were not found to functionalise cages, but catalysed enol-ketone tautomerisation in the ketone substrate.

Dodecaamide molecules could be incorporated into larger functional networks, which could incorporate cages into robust films and membranes. Furthermore, singly functionalised cages could be developed as monomers for incorporation into polymeric materials, which would enable cage cavities to be evenly spaced across a film or membrane. This could be accomplished by co-polymerisation with a non-porous monomer such as styrene, which would "pad" out the space between cage molecules. Cage polymers would represent a significant advance over previous attempts, which incorporated cage molecules *via in situ* crystallisation of covalent cages into a polymer membrane. These produced crystallites, in which the cage-cage spacing was dictated by cage-cage packing, rather than by programmable monomer to cage-monomer ratios.

5.3 Determining the molar-mass dispersity of polymers without isolation

We have developed a mathematical model for determining the dispersity of polymers in solution by analysing the PFG-NMR response curve, which gradually deviates from the linear response curve we expect for monodisperse species. By compartmentalising the response curve, we separate the signal associated with molecular size from the signal associated with dispersity. This decreases the

uncertainty produced from modelling dispersity. Resultantly, we are able to calculate dispersity by NMR to a certainty of ± 0.16 $Đ_M$.

This model was characterised for a range of polymeric species, including ideal polymers such as linear polystyrenes, and more challenging polymers such as PHEAA, which interact strongly in solution. This technique has the potential to be accurate *in situ* and was used to measure the dispersity of a linear polystyrene undergoing RAFT polymerisation in dioxane-d_8. We envisage that this could be used to characterise polymers without the need for isolation. However, significant work should be undertaken to ensure that the method is accurate, precise, robust, and applicable to a variety of solvents.

The procedure described in Chapter 4 should be followed at least three times to provide evidence that $Đ_M$ measured by NMR is precise (Table 5.1, Entries 1–3). This should be repeated at a nominal styrene concentration of 50 % and 10 % (Entries 4–7) to determine whether the dispersity is independent of concentration. Furthermore, the target degree of polymerisation should be varied between 10 % and 200 % of that previously tested (Entries 8–10). All experiments should be compared with GPC data of the same samples to assess accuracy.

The degree to which interaction between polymers in solution can affect $Đ_M$ as measured by NMR should be determined by increasing the quantity of initiator so as to increase the dispersity of the resultant polymer (Entries 11–15). It is likely that a high degree of dispersity will result in an upwards bias in the NMR measurements, as noted in Section 4.4. These experiments will determine the upper bound of accuracy for NMR measurements. As it is likely that this upper bound will depend on the concentration, this process should be repeated at a nominal styrene concentration of 50 % and 10 % (not contained in Table 5.1).

A preliminary assessment of the suitability of this technique to other polymers can be made by performing selected experiments using different monomers, such as vinyl acetate (Entries 16–19) and 2-hydroxyethyl acrylate (Entries 20–23).

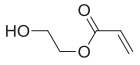
Entry	Monomer	Target DP	Styrene concentration (mg mL^{-1})	Target $Đ_M$
1		30	250	1.10
2		30	250	1.10
3		30	250	1.10
4		30	125	1.10
5		30	125	1.10
6		30	125	1.10
7		30	25	1.10
8		15	250	1.10
9		60	250	1.10
10		120	250	1.10
11		30	250	1.15
12		30	250	1.20
13		30	250	1.25
14		30	250	1.50
15		30	250	1.75
16		30	250	1.10
17		30	125	1.10
18		30	25	1.10
19		15	250	1.10
20		30	250	1.10
21		30	125	1.10
22		30	25	1.10
23		15	250	1.10

Table 5.1. A robustness study into the dispersity method described in Chapter 4.

Once accomplished, this method could be combined with continuous or stop-flow NMR techniques to facilitate large-scale reaction analysis without the need for reaction sampling.

In dilute solutions, we have shown that analysis can be conducted in a variety of solvents, and at a variety of temperatures. While solvent choice frequently affects the dispersity measured by GPC, PFG-NMR results were shown to be independent of solvent, which may indicate that a different molecular parameter is being measured. We note that whilst GPC analysis is dependent on minimising polymer-stationary-phase interactions, as well as polymer-polymer interactions, no stationary phase is present in NMR analysis. Thus, we hypothesise that PFG-NMR may present a less complicated analytical technique, which may be more reproducible across a variety of solvent types.

5.4 General conclusions

In this thesis, I have focussed on generating mathematical models that can be used to characterise structural features of molecules in solution by PFG-NMR. The aim of this has been to extend the range of techniques available during synthesis for determining the likely properties of molecules. I have demonstrated the importance of molecular shape in analysing diffusion NMR data of covalent cages, and began to investigate the range of values that could be obtained by varying the shape under analysis. This was important in determining structural features of a series of cages that alternated between prismatic and tetrahedral shape. While I have not fully characterised this new molecular descriptor, ρ_r, I have begun to establish a baseline by which the diffusion NMR of cages can be judged. Several topologies, such as prisms

and tetrahedra, have demonstrated consistent results, which could be used to justify structural assignment in the future.

I have developed a new, accessible technique for determining the molar-mass dispersity of polymers in solution without the need for isolation. Although this technique has only been validated for linear polymers, I have laid the groundwork for future studies into different classes of branched polymers, such as star, bottle-brush and crosslinked species. These are often more difficult to analyse by GPC.

Synthesising a porous polymer from porous cage monomers presents a fascinating opportunity to control the pore dimensions and spacing in solution-processable solids. While I was not successful in generating a polymeric species, I was successful in limiting the rate of reactions of cages and have produced samples rich in singly functionalised material. Although this was achieved with an isobutyric group, this forms the basis for future work in which more easily-functionalised handles could be introduced.